MAXIMUM LIKELIHOOD ESTIMATION OF DETECTOR EFFICIENCIES IN POSITRON EMISSION TOMOGRAPHY

By
WEN-HSIUNG LEE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2001

Copyright 2001

by

Wen-Hsiung Lee

ACKNOWLEDGMENTS

First I would like to thank my advisor and mentor, Professor John M. M. Anderson, for his endless effort and support on my research. His encouragement and suggestions have helped me through all the difficulties. More importantly, he taught me the methodology of solving problems, and he showed his diligence and patience in research. In short, he set a good example to me, which I believe will benefit me for all my life.

Sincere thanks go to my committee members, Dr. Edmonson, Dr. Li, and Dr. Pardalos, for their guidance and patience. Sincere thanks also go to Dr. Vatow for his constructive comments and invaluable help.

I also feel thankful toward Hong and Kerkil, my colleagues, who made me enjoy a pleasant team-working environment.

I owe a large measure of gratitude to my mother and my sisters for their unwavering support to help me go through tough times. Without them, I would not have had the strength to go on with my study at UF.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

ABSTRACT vi

CHAPTERS

1 INTRODUCTION 1
1.1 Overview of PET 2
1.2 Basic Physics of PET 4
1.3 Errors in PET 5
1.3.1 Non-Collinearity 5
1.3.2 Positron Range 5
1.3.3 Attenuation 6
1.3.4 Accidental Coincidences 6
1.3.5 Scatter 6
1.3.6 Detector Deadtime 7
1.3.7 Detector Inefficiency 7
1.4 Mathematical Model for PET 8
1.4.1 Poisson Model for Emission Scan Data 8
1.4.2 Modified PET Model for Emission Scan Data 9
1.5 Existing Methods for Estimating Detector Efficiencies 10
1.6 Outline of Research 13

2 ML ESTIMATION ALGORITHMS USING A POISSON DATA MODEL . 16
2.1 Poisson Model I 16
2.1.1 Estimation Problem Associated with Poisson Model I 21
2.1.2 Estimation Algorithms associated with Poisson Model I 21
2.2 Poisson Model II 30
2.2.1 Estimation Problem Associated with Poisson Model II 33
2.2.2 Estimation Algorithms Associated with Poisson Model II . . . 34

3 ML ESTIMATION ALGORITHMS USING A REFINED DATA MODEL . 38
3.1 Poisson Model III 38
3.1.1 Estimation Problem Associated with Poisson Model III 44
3.1.2 Estimation Algorithms Associated with Poisson Model III . . 44

4 SIMULATION 48
4.1 Synthetic Data 49
4.1.1 Comparison of Detector Efficiency Estimates 50
4.1.2 Comparison of Reconstructed Emission Images 54
4.2 Real Data 56

5 CONCLUSIONS 75

6 FUTURE WORK 77

APPENDICES

A CONCAVITY OF THE LOG-LIKELIHOOD FUNCTIONS 78
A.1 Poisson Model I 78
A.2 Poisson Model II 81
A.3 Poisson Model III 82

B CRAMER-RAO LOWER BOUNDS 84
B.1 Poisson Model I 84
B.2 Poisson Model II 86
B.3 Poisson Model III 88

C DERIVATION OF THE E STEP OF THE EM ALGORITHMS 89
C.1 Poisson Model I 89
C.2 Poisson Model II 92
C.3 Poisson Model III 94

REFERENCES 96

BIOGRAPHICAL SKETCH 100

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MAXIMUM LIKELIHOOD ESTIMATION OF DETECTOR EFFICIENCIES IN POSITRON EMISSION TOMOGRAPHY

By

Wen-Hsiung Lee

December 2001

Chairman: John M. M. Anderson
Major Department: Electrical and Computer Engineering

Positron emission tomography (PET) is a medical imaging modality that enables physicians and researchers to study biochemical process within the human body and small animals. In PET, a subject is administered a radiopharmaceutical that is absorbed preferentially by the region-of-interest (e.g., brain). Due to the radiopharmaceutical and certain interactions at the cellular level, photon pairs are emitted in all directions and the number of photon pairs emitted from a region is proportional to the metabolic rate of the region. PET scanners are designed to detect and count the photon pairs. From this photon count data, images are reconstructed whereby the value of a pixel is proportional to the metabolic rate of the region associated with the pixel. The reconstructed images provide valuable information to help physicians detect abnormalities, such as tumors, which have extremely high metabolic rates.

There are many sources of error in PET that must be addressed in order to obtain accurate images. The focus of this research is the problem of detector inefficiency. Due to inherent detector non-uniformity, detector efficiencies are less than 100%. To obtain more precise emission images, the photon count data must be corrected to account for the effect of detector inefficiency.

We introduce a maximum-likelihood method for estimating detector efficiencies in PET. First, we develop three Poisson models for blank scan data obtained from rotating rod sources. Then, for each model, we estimate the detector efficiencies using expectation-maximization algorithms, where the maximization step is solved using two optimization algorithms. As desired, the resulting algorithms have the property that the log-likelihood function is non-decreasing as the iteration number increases. For each data model, one of the proposed algorithms guarantees that the efficiency estimates always lie in the interval [0, 1]. Although the second algorithm for each data model does not have this property, in all of the experiments performed it produced efficiency estimates that were between zero and one. Simulation studies using synthetic data demonstrate that, based on various comparison criteria, the proposed estimation algorithms outperform two alternative approaches. Additionally, simulation studies using real data show that the third Poisson model leads to much better reconstructed emission images than the first two models.

CHAPTER 1
INTRODUCTION

Medical imaging modalities are widely used to obtain anatomical and biochemical information about the human body. For example, X-ray computed tomography and magnetic resonance imaging are used to acquire detailed images of the anatomical structure of a region of interest (ROI) (e.g., brain, heart, thorax). In many situations, however, it is more important to obtain functional information than anatomical information [1]. For instance, the knowledge of blood flow rates, oxygen utilization rates, and glucose utilization rates of brain or heart might be imperative for diagnosing certain diseases.

Biochemical and functional information can be obtained via a class of imaging techniques known as nuclear medicine imaging. Among the techniques within nuclear medicine imaging are planar imaging [2, 3], single-photon emission computed tomography (SPECT) [4–6], and positron emission tomography (PET) [7].

Planar imaging suffers from relatively low availability of radiopharmaceuticals, poor sensitivity of cameras, and, most importantly, the so-called projection problem. The projection problem refers to situations where the ROI might be obscured due to activity in front of or behind the ROI. Moreover, photons originating from the ROI might be attenuated due to overlying tissue.

Although SPECT overcomes the projection problem and provides more sensitivity than planar imaging, it still has insufficient sensitivity and quantitative accuracy for certain applications. Furthermore, SPECT uses radiopharmaceuticals that incorporate long half-life radionuclides, such as $^{123}\mathbf{I}$ and $^{201}\mathbf{TI}$. Because of the long half-lives

(ranging from hours to several days) of the radionuclides, patients are malignantly exposed to radioactivity for a long period of time.

PET avoids those afore-mentioned shortcomings and provides better images in the sense of efficiency, sensitivity and accuracy. In addition, positron-emitting isotopes of carbon, oxygen, nitrogen and fluorine can be readily incorporated into many useful radiopharmaceuticals because these compounds are naturally used by the body [8]. However, because radionuclides incorporated into radiopharmaceuticals for PET have short half-lives (2 minutes, 10 minutes, 20 minutes, and 110 minutes for $^{11}\mathbf{C}$, $^{15}\mathbf{O}$, $^{13}\mathbf{N}$ and $^{18}\mathbf{F}$, respectively), on-site cyclotrons are needed to produce radionuclides. The high cost and high complexity of cyclotrons have prevented PET from dominating the field of nuclear medical imaging. Even so, PET still has a wide range of use in clinical and research applications due to its advantages over other modalities.

There are a number of error sources in PET, such as attenuation, accidental coincidences, and detector inefficiency. In this dissertation, we address the problem of detector inefficiency and propose algorithms, based on the maximum likelihood estimation principle, that estimate detector efficiencies.

1.1 Overview of PET

A PET study begins with the injection or inhalation of a radiopharmaceutical by a subject. A radiopharmaceutical is a substance that is labeled with a positron-emitting isotope. The radiopharmaceutical is transported to and absorbed by the ROI. When the isotope decays, it emits positrons, which travel a short distance (on the order of tenth of millimeter) before annihilating with electrons. The areas with a high concentration of the radiopharmaceutical emit more positrons than the areas of low concentration. The annihilation of a positron with a nearby electron produces two high-energy (511 keV) photons that move away from each other in nearly opposite directions. The position where the annihilation takes place is unknown and

the photons' line of flight is omni-directional and completely random. If these two photons, after traveling through the subject, are detected by a pair of detectors on a detector ring surrounding the ROI within a small timing window τ ($\sim$ 10ns), a *coincidence* is recorded by that detector pair or tube (a detector pair is also referred to as a *tube*.) The observed *emission data* are the collection of the counts (i.e., total number of coincidences) recorded by each detector tube. An illustration of the scanning process is provided in Fig. 1.1, which shows a cross-section of the scanner.

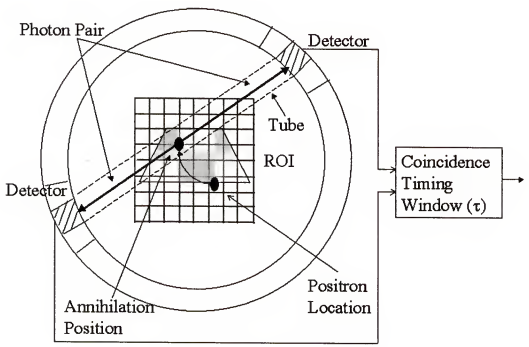


Figure 1.1: Annihilation of a positron with a nearby electron produces a pair of photons that propagate in nearly opposite directions.

The number of photons emitted from a region is proportional to the concentration of radiopharmaceutical in the region. Moreover, the concentration of the radiopharmaceutical in the region is proportional to the metabolic rate of the region. Using the emission data, reconstruction algorithms create images, called *emission images*, where the value of a pixel is proportional to the concentration of the radiopharma-

ceutical in the region associated with the pixel. Thus, emission images can be used to assess biochemical processes.

1.2 Basic Physics of PET

The cylindrical rings of detectors surrounding the ROI typically have a diameter of 80-100 cm and an axial extent of 10-20 cm. The detectors are shielded from outside the field of view (or ROI) by relatively thick, lead end-shields. Most scanners can switch between two operating modes. In the first mode, known as slice-collimated mode or 2-dimensional (2D) mode, axial collimation is provided by thin annular rings of tungsten called *septa*, and only *direct plane* sinograms involving coincidences between detectors on the same ring are recorded. In practice, the situation may be more complicated in that, in order to improve statistics, direct and cross plane sinograms can be combined in different ways [9]. The other mode is fully three-dimensional (3D) mode where septa are retracted and *cross plane* sinograms of coincidences between detectors on different rings are also recorded [7].

Detectors are critical components of a PET scanner [7]. Each detector ring usually contains several hundred detector modules, where each module consists of an array of crystals coupled to several photomultiplier tubes. When a photon interacts with a crystal, electrons are moved from the valence band to the conduction band. Then, the electrons return to the valence band at impurities in the crystal. This process causes numerous light photons to be emitted. A photomultiplier tube collects light photons and converts them into electrical signals, which are decoded so as to determine which detector module was actually hit by a photon. The material bismuth-germanate is widely used as the crystal because of its high attenuation density, high light output (2500 light photons per 511 keV photon), fast rise time, and short decay time (300 ns). Crystals with high attenuation densities are desirable because the probability of a photon interaction is very high. Moreover, a high light output improves positioning

accuracy, while a fast rise time is helpful for accurate timing, and a short decay time enables the scanner to handle high counting rates.

1.3 Errors in PET

There are two major categories of physical effects in PET: (1) effects due to the positron annihilation process and propagation of photons through the body, such as positron range, attenuation, Compton scatter, and accidental coincidences, and (2) effects that result from the interactions of photons with detectors, such as detector efficiency and dead-time [7].

1.3.1 Non-Collinearity

As mentioned previously, an emitted positron interacts with an electron as it travels through the body. Each interaction requires certain energy. When the momentum of the positron's energy is nearly zero, annihilation occurs and two annihilation photons are produced each with an energy of 511 keV. The photon pair flies off along nearly collinear paths, where the degree of non-collinearity depends on the energy left in the positron [7]. The energy of the positron after annihilation varies widely among isotopes. The divergence from collinearity is on the order of a degree or less. So it is usually ignored.

1.3.2 Positron Range

The distance the positron travels before annihilation is termed *positron range.* Typically, the positron range is about 3-5 μm [7]. Since the resolution of reconstructed PET images is over 5 mm for most commercial scanners, the error introduced by positron range is not a serious problem. However, for high resolution experimental scanners, this becomes the fundamental limitation of PET.

1.3.3 Attenuation

Two interactions that may cause a photon pair to go undetected are photoelectric absorption and Compton scattering. For 511 keV photons, the incidence of photoelectric absorption is very low and can be ignored in most cases [7]. A Compton scatter happens when a photon interacts with an outer shell electron. In this case, the photon is deflected from its path. Consequently, the counts of the tube corresponding to the original path of a photon pair does not increase when one or both of the photons undergo Compton scattering. The total effect of photoelectric absorption and Compton scatter is called *attenuation* and undetected annihilations are termed *scattered* events.

1.3.4 Accidental Coincidences

If two photons resulting from different annihilations are detected within the coincidence timing window, they will be incorrectly interpreted as a coincidence event. Such events are called *accidental coincidences* or *randoms.* as *true coincidences* or *trues.*

The rate of accidental coincidences is related to the concentration of the radionuclide in the ROI. A common correction strategy is to subtract an estimate of the randoms by using an estimate of the single events or delayed coincidence circuitry [10].

1.3.5 Scatter

Suppose one or both photons of a photon pair undergo Compton scattering. If the photons remain in the detector plane and are detected, then the annihilation event is recorded by the wrong tube. In other words, since the photons' path is not collinear, the event is mis-positioned and additive noise results. However, due to the energy loss caused by the scattering process, scattered photons can be distinguished from

unscattered photons to a limited degree by setting a threshold such that photons with sufficiently low energy, which correspond to scattered photons. are ignored.

1.3.6 Detector Deadtime

The time required to process a single photon event limits the counting rate of a PET scanner [11]. Event processing begins with the rising edge of the pulse for the first detector involved. The pulse is integrated for some time interval, after which position calculations and energy discriminations are performed. The detectors are "dead" to new photon events during this time. Since the number of randoms increase as the square of the activity in the field of view [7], the number of detected counts may decrease as the flux of incident photons increases. This effect is accounted for in two ways. First, it becomes a dominant limitation on the injected dose. Second, a scale factor is used to correct data for the deadtime.

1.3.7 Detector Inefficiency

The efficiency of a detector pair is the probability that an incident photon pair is detected by the detector pair. Due to inherent detector non-uniformities, detector efficiencies are not equal and less than 100% [12, 13]. The efficiency of current PET detectors depends on a number of geometric and non-geometric factors. The non-geometric factors include fluctuations in photomultiplier tube gains and variations of crystal characteristics [7]. Geometric factors include the angles subtended by detectors and the angles of incidence of photons. The angle subtended by a detector pair depends only on the perpendicular distance from the origin, and the angle subtended by the detector pair affects the angles of incidence of photons. And the length of intersection of a detector and the path of a photon depends on the angle of incidence of the photon. If the intersection length is larger, the probability that the photon is detected by the detector is higher.

1.4 Mathematical Model for PET

Since the data in PET follow from a counting process, it is assumed that the emission process is a spatial Poisson process. In 1982, Shepp and Vardi [14] proposed a *Poisson model* for PET that has been widely accepted as a valid mathematical model for the PET process. However, their model was based on the assumption that the data are error-free except for Poisson noise. In the following, the basic model for PET emission data is first introduced. Because this dissertation focuses on how to correct (or normalize) emission data for variations in detector efficiency, the basic data model is then modified to take into account the effect of detector inefficiency.

1.4.1 Poisson Model for Emission Scan Data

The ROI is divided into a grid of J voxels and the emission intensity is assumed to be constant over each voxel. Assuming the data are free of errors, the emission datum recorded by detector pair (k,l) is a realization of a random variable Y_{kl}:

$$Y_{kl} \sim \text{Poisson}([\boldsymbol{P}\boldsymbol{x}]_{kl}), \quad k = 1, 2, \ldots, I, \; l \in \text{fan}_k,$$

where

$$[\boldsymbol{P}\boldsymbol{x}]_{kl} \triangleq \sum_{j=1}^{J} P_{kl,j} x_j, \tag{1.1}$$

and where I is the number of detectors on a ring and the set fan_k is the set of detectors that are paired with detector k (see Fig. 1.2). In addition, $P_{kl,j}$ is the probability that a photon pair emitted from voxel j is incident on detector pair (k,l), x_j is the mean number of photon pairs emitted from voxel j, and $\boldsymbol{x} \triangleq [x_1, x_2, \ldots, x_J]^T$ is the *emission intensity vector*.

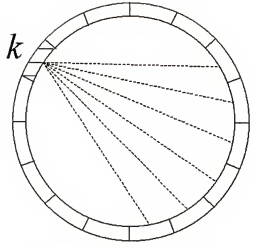


Figure 1.2: Definition of fan_k.

1.4.2 Modified PET Model for Emission Scan Data

Again, the efficiency of a detector pair is the probability that an incident photon pair is detected by the detector pair. Let ϵ_{kl} be the efficiency of detector pair (k,l). Having a detector pair that is not 100% efficient is mathematically equivalent to *thinning* [15] the incident photons. To see what effect the thinning has on a Poisson random variable, we make use of the following well-known Corollary.

Corollary *Let Z be a Poisson random variable with mean $\bar{Z}$. Suppose a random variable Y is created by thinning Z with probability p. Then, Y is Poisson with mean $p\bar{Z}$.*

Proof:

$$\begin{aligned} P(Y=y) &= \sum_{z=y}^{\infty} P(Y=y \mid Z=z)P(Z=z) \\ &= \sum_{z=y}^{\infty} \binom{z}{y} p^y (1-p)^{z-y} \cdot e^{-\bar{Z}} \frac{(\bar{Z})^z}{z!} \\ &= \sum_{z=y}^{\infty} \frac{z!}{y!(z-y)!} p^y (1-p)^{z-y} \cdot e^{-\bar{Z}} \frac{(\bar{Z})^z}{z!} \end{aligned}$$

$$= e^{-\bar{z}} \frac{p^y}{y!} \sum_{z=y}^{\infty} \frac{(1-p)^{z-y}(\bar{Z})^z}{(z-y)!}.$$

Using the change of variable $z' = z - y$, the probability $P(Y = y)$ becomes

$$\begin{aligned} P(Y=y) &= e^{-\bar{z}} \frac{p^y}{y!} \sum_{z'=0}^{\infty} \frac{(1-p)^{z'}(\bar{Z})^{z'+y}}{z'!} \\ &= e^{-\bar{z}} \frac{(p\bar{Z})^y}{y!} \sum_{z'=0}^{\infty} \frac{((1-p)\bar{Z})^{z'}}{z'!} \\ &= e^{-\bar{z}} \frac{(p\bar{Z})^y}{y!} e^{(1-p)\bar{Z}} \\ &= \frac{e^{-p\bar{Z}}(p\bar{Z})^y}{y!}. \end{aligned} \tag{1.2}$$

Thus, Y is a Poisson random variable with mean $p\bar{Z}$.

Q.E.D.

According to the above Corollary, the emission data, after thinning with probability ϵ_{kl}, become

$$Y_{kl} \sim \text{Poisson}(\epsilon_{kl}[\boldsymbol{P}\boldsymbol{x}]_{kl}), \quad k = 1, 2, \ldots, I, \ l \in \text{fan}_k. \tag{1.3}$$

Given the observed emission scan data $\boldsymbol{y}$, which is a vector with elements $\{y_{kl}\}$, the problem is to reconstruct the emission intensity vector $\boldsymbol{x}$. Shepp and Vardi [14], the developers of the Poisson model for emission scan data, assumed that the detectors were 100% efficient. However, detector efficiencies must be estimated before emission images can be reconstructed, because, in reality, they are not ideal.

1.5 Existing Methods for Estimating Detector Efficiencies

To estimate a set of detector efficiencies, a scan is taken using either rotating rods or a uniform cylinder that contains positron emitting nuclides. The scan, which is referred to as a *blank scan* (no patient in the scanner), lasts about 2-4 hours and is usually performed once a month (see Fig. 1.3).

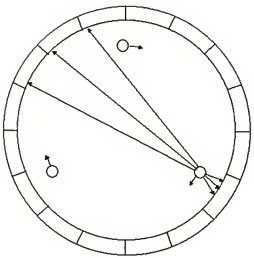


Figure 1.3: Blank scan.

Numerous methods [9, 16–24] have been developed for estimating detector efficiencies. Unfortunately, existing methods for estimating the detector efficiencies are largely ad hoc and do not take into account the statistical properties of the blank scan data. For instance, an early technique discussed in [9] was to estimate the efficiency of a detector pair by computing the ratio of the number of photon pairs measured by that detector pair and the average number of photon counts per tube. It has been noted that the accuracy of this heuristic technique is limited by the statistical accuracy of the blank scan [9]. Another technique for estimating detector efficiencies is the well known *fan-sum* method by Hoffman et al. [16]. Hoffman et al. assumed that b_{kl}, the number of photon pairs recorded by detector tube (k, l), is equal to

$$b_{kl} = \epsilon_k \epsilon_l n_{kl}, \quad k = 1, 2, \ldots, I, \; l \in \text{fan}_k, \tag{1.4}$$

where ϵ_k is the efficiency of the k^{th} detector and n_{kl} is the unknown number of photon pairs *transmitted* along detector tube (k, l). Hoffman et al. also assumed that

$$\sum_{l \in \text{fan}_k} \epsilon_l n_{kl} \approx \text{constant}. \tag{1.5}$$

Their reason for making the assumption in (1.5) is because the summation is over a large number of detectors and the variations would tend to average out. Under these two assumptions, the proposed estimate for the efficiency of the k^{th} detector

$$\hat{\epsilon}_k = \sum_{l \in \text{fan}_k} b_{kl} \tag{1.6}$$

is approximately equal to

$$\hat{\epsilon}_k \approx \epsilon_k \cdot \text{constant}. \tag{1.7}$$

The approximation in (1.5) is poor when there are large fluctuations in the detector efficiencies [17] or the number of terms in the set fan_k depends on k. In such instances, the fan-sum algorithm may not provide accurate estimates for the detector efficiencies. Also, the fan-sum algorithm does not account for the Poisson nature of the blank scan data.

Several methods [9,16–19] have been developed based on the assumption of noise-free (i.e., no Poisson noise) blank scan data. For example, an iterative method by Ferreira et al. [17] is based on the assumptions that the data are noise free and every detector views the same activity:

$$\sum_{l \in \text{fan}_k} n_{kl} = A, \quad k = 1, 2, \ldots, I, \; l \in \text{fan}_k. \tag{1.8}$$

Substituting (1.4) into (1.8) gives

$$\epsilon_k = \frac{1}{A} \sum_{l \in \text{fan}_k} \frac{b_{kl}}{\epsilon_l}, \quad k = 1, 2, \ldots, I, \; l \in \text{fan}_k. \tag{1.9}$$

To get an approximate solution for the above system of equations, the following fixed-point iteration is used:

$$\epsilon_k^{m+1} = \sum_{l \in \text{fan}_k} \frac{b_{kl}}{\epsilon_l^m}, \quad k = 1, 2, \ldots, I, \; l \in \text{fan}_k, \tag{1.10}$$

where ϵ_k^m is the m^{th} estimate of ϵ_l. The constant factor $1/A$ is neglected because the efficiency estimates are normalized so that their mean is one. Like the fan-sum algorithm, Ferreira's algorithm assumes noise-free data and relies on a key assumption (see (1.8)) that is questionable when the number of terms in the set fan_k depends on k.

Another algorithm based on the assumption of noise-free data is the algorithm by Defrise et al. [9]. The algorithm by Defrise et al. parallels the fan-sum method with the exception that geometric means rather than arithmetic means of the data are used to estimate the mean number of unobserved number of transmitted photon pairs. Defrise et al. claimed that their algorithm produces unbiased efficiency estimates in the sense that they converge to the real efficiencies in the ideal case where the blank scan data are noise-free.

1.6 Outline of Research

All of the estimation methods we have discussed are sub-optimal in the sense that they assume noise free data and do not provide ML estimates. To capture the stochastic nature of PET data, we propose three Poisson models for blank scan data. At first we ignore the problem of detector penetration [25–28] when developing the

first two Poisson models for blank scan data. However, later we incorporate the effect of detector penetration and refine our original Poisson models for blank scan data.

Motivated by the optimality properties of ML estimators [29], we maximize the log-likelihood function associated with each Poisson model by using the expectation-maximization (EM) algorithm [30, 31]. To apply the EM algorithm, we choose the blank scan data and unobserved number of transmitted photon pairs for each detector tube to be the complete data. Two possibilities have been explored:

- Over-parameterized approach, where the unknowns are the efficiencies of the detector pairs, $\{\epsilon_{kl}\}$, and certain emission means.
- Sufficiently-parameterized approach, where unknowns are the efficiencies of individual detectors, $\{\epsilon_k\}$, and certain emission means.

Because the over-parameterized approach does not take into consideration the assumption that $\epsilon_{kl} = \epsilon_k \epsilon_l$, $k = 1, 2, \ldots, I$, $l \in \text{fan}_k$, the number of unknowns is much higher than that for the sufficiently-parameterized approach. Closed form expressions for the E step and M step of the EM algorithm are derived for our first model using the over-parameterized formulation. However, simulations show that the over-parameterized approach has poor performance. Thus, the results for the over-parameterized approach will not be presented in Chapter 4, which contains our simulation results.

For each of our three proposed Poisson models for blank scan data, we use the sufficiently-parameterized formulation and develop two algorithms for estimating detector efficiencies. For each data model, a closed form expression for the E step of the EM algorithm is derived. Unfortunately, no closed form solutions to the maximization problems in the M step are available. For each data model, we propose solving the maximization problems in the M step iteratively by using a fixed-point iteration or the method of coordinate descent. All the resulting (EM) algorithms have

the property that the log-likelihood function increases monotonically with increasing iteration. It is desirable that a detector efficiency estimator provides estimates that lie in the interval [0, 1]. The coordinate descent based algorithms we propose for each data model guarantee that the efficiency estimates always lie in the interval [0, 1]. The fixed-point iteration based algorithms for each data model do not have this property. However, in all the experiments we performed the efficiency estimates did lie in the interval [0, 1]. In terms of certain objective criteria, simulation studies using synthetic data demonstrate that the proposed estimation algorithms outperform the fan-sum and Ferreira's algorithms. Additionally, our experiments show that emission images based on the proposed detector efficiency estimates have smaller mean squared error than emission images obtained using the detector efficiency estimates generated by the fan-sum and Ferreira's algorithms. Simulation studies using real data show that the estimation algorithms based on the third Poisson model provide much better reconstructed emission images.

This dissertation is organized as follows. In Chapter 2, we introduce two Poisson data models for blank scan data and develop estimation algorithms that provide ML estimates of the detector efficiencies. Then, in Chapter 3, we refine our data model to incorporate the effect of detector penetration and present the corresponding ML estimation algorithms. In Chapter 4, we discuss the results of our simulation studies. In Chapter 5, conclusions about our research are given. Finally, we suggest some topics for future work in Chapter 6.

CHAPTER 2
ML ESTIMATION ALGORITHMS USING A POISSON DATA MODEL

In this chapter, we first present two Poisson models for blank scan data obtained from rotating rod sources. Then, we present ML estimation algorithms based on these two Poisson models. The first model is based on the fact that the mean number of photon pairs transmitted along a detector pair only depends on its perpendicular distance from the origin, which is the center of field-of-view. The second model, which is a modification of the first model, takes into consideration that the rods are nearly uniform. Note, the second model has fewer unknowns than the first model.

In many scanners, the space between blocks of detectors is filled up with lead. Throughout our research, though, we make the simplifying assumption that there is no lead between blocks of detectors. Also, we assume the detector rings are perfectly circular.

2.1 Poisson Model I

Before discussing the first Poisson model, we introduce some notation. Let α denote the angle between the lines that connect the origin to detectors k and l, and $p(k,l)$ denote the perpendicular distance from the origin to the line defined by the mid-points of detectors k and l (see Fig. 2.1). The detector indices k and l are defined such that $k < l$, and the angle α is defined counterclockwise starting from detector k and ending at detector l. Clearly, $p(k,l) = r\cos(\frac{\alpha}{2})$, where $\alpha = \frac{2\pi}{I}(l-k)$ and r is the radius of the detector ring. Therefore, for detector pairs (k_1,l_1) and (k_2,l_2), $p(k_1,l_1) = p(k_2,l_2)$ if $l_1 - k_1 = l_2 - k_2$. In addition, $p(k_1,l_1)$ and $p(k_2,l_2)$ will

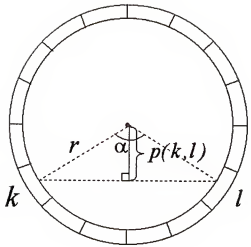


Figure 2.1: Definitions of $p(k,l)$ and α.

be equal if $(l_1 - k_1) + (l_2 - k_2)$ equals the number of detectors on one ring. To see this fact, consider Fig. 2.2, where detector pairs (1,114) and (50,321) of the Siemens-CTI ECAT EXACT 921 scanner [26] are depicted (note: this scanner has $I = 384$ detectors per ring). Because $\alpha_1 = \frac{2\pi}{384}(114 - 1)$ and $\alpha_2 = \frac{2\pi}{384}(321 - 50)$, their sum

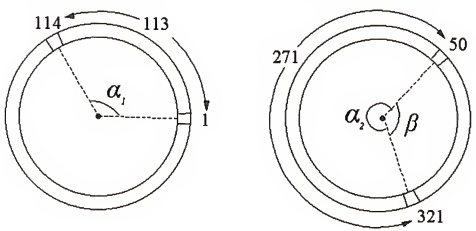


Figure 2.2: An example to demonstrate how $p(k,l)$ is related to the value of $l - k$.

equals $\alpha_1 + \alpha_2 = \frac{2\pi}{384}(113 + 271) = 2\pi$. Clearly $\alpha_2 + \beta = 2\pi$ as well. Therefore, $\alpha_1 = \beta$, which implies $p(1,114) = p(50,321)$.

For the Siemens-CTI ECAT EXACT 921 scanner, Table 2.1 illustrates the detector pairs of the first projection and the perpendicular distance of each pair. From the

Table 2.1: Perpendicular distances of detector pairs for the first projection of the Siemens-CTI ECAT EXACT 921 scanner.

1st projection			
member	pair (k,l)	$l-k$	$p(k,l)(cm)$
1st	(152,264)	112	$d_1 = 25.1114$
2nd	(152,265)	113	$d_2 = 24.8428$
3rd	(151,265)	114	$d_3 = 24.5726$
⋮	⋮	⋮	⋮
80th	(113,304)	191	$d_{80} = 0.3375$
81st	(112,304)	192	$d_{81} = 0$
82nd	(112,305)	193	$d_{80} = 0.3375$
⋮	⋮	⋮	⋮
159th	(73,343)	270	$d_3 = 24.5726$
160th	(73,344)	271	$d_2 = 24.8428$

discussion above, it can be seen that the value of $l-k$ determines the perpendicular distance $p(k,l)$. Taking this property into account, it can be seen that there are 81 possible perpendicular distances from the origin.

To see how the number of transmitted photon pairs is related to $p(k,l)$, consider a scanner with 16 detectors, where the dashed circle stands for the path of the rods (see Fig. 2.3). It is easy to show that the arc of tube 2 is longer than the arc of tube 1. Thus, because the rods rotate at a constant speed, the rods will stay in tube 2 for a longer time than in tube 1. Observe that $p(k,l)$ of tube 2 is larger than $p(k,l)$ of tube 1. More generally, it can be proved using a simple geometric argument that the dwell time is proportional to the perpendicular distance from the center to the tube. And, because the activity viewed by a detector tube is proportional to the dwell time of the rods in that tube, it is reasonable to claim that the mean number of the transmitted photon pairs depends only on the perpendicular distance from the origin to the tube. For a blank scan, let B_{kl} denote the observed number of photon pairs *detected* by detector pair (k,l), and N_{kl} denote the unobserved number of *transmitted* photon pairs along detector pair (k,l). Given the discussion above, we assume that

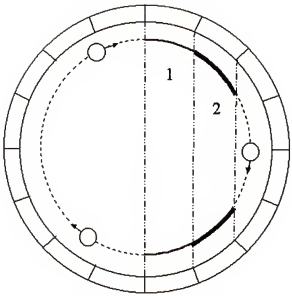


Figure 2.3: An example to show how the number of transmitted photon pairs depends on $p(k,l)$.

B_{kl} and N_{kl} are Poisson with means $\epsilon_{kl}\lambda_{p(k,l)}$ and $\lambda_{p(k,l)}$, respectively:

$$N_{kl} \sim \text{Poisson}(\lambda_{p(k,l)}),\ k = 1,2,\ldots,I,\ l \in \text{fan}_k, \tag{2.1}$$

$$B_{kl} \sim \text{Poisson}(\epsilon_{kl}\lambda_{p(k,l)}),\ k = 1,2,\ldots,I,\ l \in \text{fan}_k, \tag{2.2}$$

where $\lambda_{p(k,l)}$ depends only on $p(k,l)$. We refer to the data model described by equations (2.1) and (2.2) as Poisson Model I. To see the suitability of Poisson Model I, we compare real blank scan data and synthetic data generated according to Poisson Model I in Figs. 2.4 and 2.5. It can be seen from the figures that the synthetic data bears a close resemblance to the real data.

Note, a rod passes through the same detector tube twice in each cycle. This fact is implicitly accounted for through the parameters $\{\lambda_{p(k,l)}\}$.

From Poisson Model I, it follows that

$$E[B_{kl}] = \epsilon_{kl}E[N_{kl}]. \tag{2.3}$$

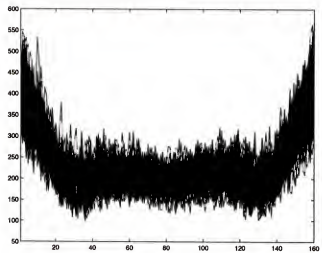


Figure 2.4: Overlapped 192 projections of real blank scan data.

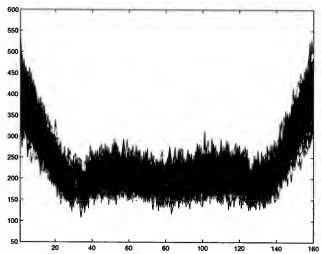


Figure 2.5: Overlapped 192 projections of synthetic blank scan data generated according to Poisson Model I.

Hoffman's assumption in (1.4) follows by replacing the expectations $E[B_{kl}]$ and $E[N_{kl}]$ in (2.3) with their ML estimates b_{kl} and n_{kl}, respectively. The statement in (2.3) is more sound probabilistically than the assumption in (1.4) because detector efficiencies should be interpreted as probabilities. Moreover, all we can say about B_{kl} and N_{kl} is that their means obey the relationship in (2.3).

2.1.1 Estimation Problem Associated with Poisson Model I

For convenience, let $\boldsymbol{B}$, $\boldsymbol{N}$, $\boldsymbol{\epsilon}$, and $\boldsymbol{\lambda}$ be vectors with elements $\{B_{kl}\}$, $\{N_{kl}\}$, $\{\epsilon_k\}$, and $\{\lambda_{p(k,l)}\}$, respectively. Also, let $\boldsymbol{b}$ be a vector with elements $\{b_{kl}\}$, where b_{kl} is an observation of B_{kl}. To estimate $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$, we use the ML method:

$$(\mathbf{P1}) \qquad (\hat{\boldsymbol{\epsilon}}, \hat{\boldsymbol{\lambda}}) = \arg\max_{\boldsymbol{\epsilon},\boldsymbol{\lambda}} \phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda}) \text{ subject to } \mathbf{0} < \boldsymbol{\epsilon} \leq \mathbf{1} \text{ and } \boldsymbol{\lambda} > \mathbf{0}, \tag{2.4}$$

where $\phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda})$ is the log-likelihood function

$$\phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda}) \triangleq \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) = \log \prod_{k=1}^{I} \prod_{l \in \text{fan}_k} e^{-\epsilon_{kl}\lambda_{p(k,l)}} \frac{(\epsilon_{kl}\lambda_{p(k,l)})^{b_{kl}}}{b_{kl}!}, \tag{2.5}$$

and $\mathbf{0}$ and $\mathbf{1}$ are $I \times 1$ vectors of zeros and ones, respectively. In Appendix A, we show that Hessian of the log-likelihood function is not negative definite. Thus, the log-likelihood function is not concave. Given this fact, it is expected that the performance of iterative algorithms used to solve (**P1**) will depend on the initialization procedure.

Cramer-Rao Lower Bound

To compute the Cramer-Rao lower bound (CRB) for the estimators of $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$ in (**P1**), we need to first compute the Fisher information matrix. We prove in Appendix B that the Fisher information matrix is non-invertible for the probability density function associated with (**P1**). Therefore, the CRB for the estimators of $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$ cannot be determined.

2.1.2 Estimation Algorithms associated with Poisson Model I

If $\boldsymbol{n}$, an observation of $\boldsymbol{N}$, was available, then the parameters $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$ could be estimated using

$$\hat{\epsilon}_{kl} = \frac{b_{kl}}{n_{kl}}, \quad k = 1, 2, \ldots, I, \; l \in \text{fan}_k, \tag{2.6}$$

$$\hat{\lambda}_d = \frac{\sum_{\{(k,l):p(k,l)=d\}} n_{kl}}{\text{number of pairs } (k,l) \text{ such that } p(k,l)=d}, \quad d = d_1, d_2, \ldots, d_D, \tag{2.7}$$

where $\{d_j\}$ are the possible perpendicular distances and D is the number of possible perpendicular distances. Given this observation, our choice for the complete data is $\boldsymbol{X} = [\boldsymbol{B}, \boldsymbol{N}]$. It should be noted that an observation of $\boldsymbol{N}$ is *not* available. Consequently, as part of the EM algorithm, an estimate of an observation of $\boldsymbol{N}$ will be generated.

The E step of the EM algorithm requires the computation of the conditional expectation of the log-likelihood function of the complete data given the incomplete data and current estimates of $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$. Let $U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m)$ denote this conditional expectation, where $\boldsymbol{\epsilon}^m$ and $\boldsymbol{\lambda}^m$ are the estimates of the detector efficiencies and mean number of transmitted photon pairs, respectively, at the m^{th} iteration. Also, let $L_c(\boldsymbol{b}, \boldsymbol{n}; \boldsymbol{\epsilon}, \boldsymbol{\lambda})$ denote the log-likelihood function of the complete data:

$$L_c(\boldsymbol{b}, \boldsymbol{n}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) \triangleq \log P(\boldsymbol{B} = \boldsymbol{b}, \boldsymbol{N} = \boldsymbol{n}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}). \tag{2.8}$$

In Appendix C, we show that

$$\begin{aligned} U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m) &\triangleq E[L_c(\boldsymbol{B}, \boldsymbol{N}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m] \\ &= \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [b_{kl} \log(\epsilon_k \epsilon_l) + (n_{kl}^m - b_{kl}) \log(1 - \epsilon_k \epsilon_l)] \\ &+ \sum_{d=d_1}^{d_D} \sum_{\{(k,l):p(k,l)=d\}} [-\lambda_d + n_{kl}^m \log \lambda_d] + C_1, \end{aligned} \tag{2.9}$$

where C_1 is a constant that includes all of the terms that are independent of $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$, and n_{kl}^m is the conditional expectation of N_{kl} given the incomplete data and the estimates $\boldsymbol{\epsilon}^m$ and $\boldsymbol{\lambda}^m$:

$$n_{kl}^m \triangleq E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m], \quad k = 1, 2, \ldots, I, \; l \in \text{fan}_k. \tag{2.10}$$

We also show in Appendix C that

$$n_{kl}^m = (1 - \epsilon_k^m \epsilon_l^m)\lambda_{p(k,l)}^m + b_{kl}, \quad k = 1, 2, \ldots, I, \; l \in \text{fan}_k. \tag{2.11}$$

The quantity n_{kl}^m is an estimate of the number of photon pairs transmitted along detector pair (k, l). From (2.11), it can be seen that n_{kl}^m is estimated by adding an estimate of the number of photon pairs that are transmitted but not detected, $(1 - \epsilon_k^m \epsilon_l^m)\lambda_{p(k,l)}^m$, to the number of photon pairs that are transmitted and detected, b_{kl}.

We recognize from (2.9) that $U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m)$ is *decoupled* in the sense that it can be written as the sum of functions f_1 and g that depend only on $\boldsymbol{\lambda}$ and $\boldsymbol{\epsilon}$, respectively. More specifically,

$$U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m) = f_1^{(m)}(\boldsymbol{\lambda}) + g^{(m)}(\boldsymbol{\epsilon}) + C_1, \tag{2.12}$$

where

$$f_1^{(m)}(\boldsymbol{\lambda}) \triangleq \sum_{d=d_1}^{d_D} \sum_{\{(k,l):p(k,l)=d\}} [-\lambda_d + n_{kl}^m \log \lambda_d], \tag{2.13}$$

and

$$g^{(m)}(\boldsymbol{\epsilon}) \triangleq \sum_{k=1}^{I} \sum_{l \in \text{fan}\, k} [b_{kl} \log(\epsilon_k \epsilon_l) + (n_{kl}^m - b_{kl}) \log(1 - \epsilon_k \epsilon_l)]. \tag{2.14}$$

From (2.12), it follows that the problems of maximizing $U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m \boldsymbol{\lambda}^m)$ with respect to $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$ are decoupled.

Summarizing our results, the steps of the EM algorithm are:

Step I. (*Initialization*)

Set m=0.

Choose $\boldsymbol{\epsilon}^0$ *and* $\boldsymbol{\lambda}^0$ *s.t.* $\mathbf{0} < \boldsymbol{\epsilon}^0 \leq \mathbf{1}$ *and* $\boldsymbol{\lambda}^0 > \mathbf{0}$.

Step II. (*E step*)

$$n_{kl}^m = (1 - \epsilon_k^m \epsilon_l^m)\lambda_{p(k,l)}^m + b_{kl}, \quad k = 1, 2, \ldots, I, \ l \in fan_k.$$

Step III. (*M step*)

$(\mathbf{P}_{\boldsymbol{\lambda}})$ $\boldsymbol{\lambda}^{m+1} \triangleq \arg\max_{\boldsymbol{\lambda}>\mathbf{0}} f_1^{(m)}(\boldsymbol{\lambda})$.

$(\mathbf{P}_{\boldsymbol{\epsilon}})$ $\boldsymbol{\epsilon}^{m+1} \triangleq \arg\max_{\mathbf{0}<\boldsymbol{\epsilon}\leq\mathbf{1}} g^{(m)}(\boldsymbol{\epsilon})$.

Step IV.

If stopping criterion is not satisfied, set $m = m + 1$ *and go to Step II.*

Otherwise stop.

Solving $(\mathbf{P}_{\boldsymbol{\lambda}})$, which is equivalent to maximizing $U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m \boldsymbol{\lambda}^m)$ with respect to $\boldsymbol{\lambda}$, is straightforward and leads to a closed-form solution. Momentarily ignoring the constraints

$$\lambda_d > 0, \quad d = d_1, d_2, \ldots, d_D, \tag{2.15}$$

we take the partial derivative of $f_1^{(m)}(\boldsymbol{\lambda})$ with respect to each λ_d and set it to zero:

$$\frac{\partial f_1^{(m)}(\boldsymbol{\lambda})}{\partial \lambda_d} = \sum_{\{(k,l):p(k,l)=d\}} (-1 + n_{kl}^m \frac{1}{\lambda_d}) = 0, \quad d = d_1, d_2, \ldots, d_D. \tag{2.16}$$

Solving (2.16) yields the desired update

$$\lambda_d^{m+1} = \frac{\sum_{\{(k,l):p(k,l)=d\}} n_{kl}^m}{\text{number of pairs } (k,l) \text{ such that } p(k,l) = d}, \quad d = d_1, d_2, \ldots, d_D. \tag{2.17}$$

From (2.11), it can be seen that the constraints in (2.15) are satisfied if $\epsilon_k^m \in [0, 1]$ and $\lambda_d^m > 0$ for all m, k and d.

Unfortunately, a closed-form solution is not available for the maximizer of the function $U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m \boldsymbol{\lambda}^m)$ with respect to $\boldsymbol{\epsilon}$ (see $(\mathbf{P}_{\boldsymbol{\epsilon}})$). We *iteratively* solve this maximization problem using either the coordinate descent (CD) method [32, pages 283-287] or a fixed-point iteration. See Abatzoglou et al. [33] or Luo et al. [34] for information on the convergence properties of CD method.

EM-CD Algorithm

To solve $(\mathbf{P}_{\boldsymbol{\epsilon}})$, we first use the CD method where the function $g^{(m)}$ is minimized iteratively on a coordinate by coordinate basis. More specifically, the desired minimizer, $\boldsymbol{\epsilon}^{m+1}$, is given by solving the following maximization problems iteratively for $k = 1, 2, \ldots, I$:

$$\epsilon_k^{m+1,i+1} = \arg \max_{0<\epsilon\leq 1} g^{(m)}(\epsilon_1^{m+1,i+1}, \epsilon_2^{m+1,i+1}, \ldots, \epsilon_{k-1}^{m+1,i+1}, \epsilon, \epsilon_{k+1}^{m+1,i}, \ldots, \epsilon_I^{m+1,i}), \quad (2.18)$$

where $\epsilon_k^{m+1,i+1}$ is the estimate of ϵ_k^{m+1} at the $(i+1)^{th}$ iteration. A reasonable choice for the initial estimate of $\boldsymbol{\epsilon}^{m+1}$ is $\boldsymbol{\epsilon}^{m+1,0} \triangleq \boldsymbol{\epsilon}^m$. The iteration is terminated when successive estimates are virtually the same. In practice, the algorithm is terminated when

$$\sum_{k=1}^{I} \left(\frac{\epsilon_k^{m+1,i+1} - \epsilon_k^{m+1,i}}{\epsilon_k^{m+1,i}}\right)^2 < \delta_\epsilon, \quad (2.19)$$

where δ_ϵ is a small positive number. The iterate that results when the stopping criterion is satisfied is considered the solution to $(\mathbf{P}_{\boldsymbol{\epsilon}})$.

To see how the maximization problems in (2.18) can be solved, let

$$\phi_l(\epsilon) \triangleq g^{(m)}(z_1, z_2, \ldots, z_{l-1}, \epsilon, z_{l+1}, \ldots, z_I), \quad l = 1, 2, \ldots, I, \quad (2.20)$$

where $\boldsymbol{z}$ is a vector whose elements, $z_1, z_2, \ldots, z_I$, all lie in the interval $[0, 1]$. It is straightforward to see that

$$\frac{\partial^2 \phi_l(\epsilon)}{\partial \epsilon^2} = \sum_{k \in \mathrm{fan}_l} \left[\frac{-b_{kl}}{\epsilon^2} - (n_{kl}^m - b_{kl}) \frac{z_k^2}{(1 - z_k \epsilon)^2}\right] \quad (2.21)$$

is less than zero, for $l = 1, 2, \ldots, I$, if $\epsilon_k^m \in [0, 1]$ and $\lambda_d^m > 0$ for all m, k and d. Hence, ϕ_l is a strictly concave function if $\epsilon_k^m \in [0, 1]$ and $\lambda_d^m > 0$ for all m, k and d. Because of this concavity property, the problems in (2.18) can be readily solved with arbitrary

accuracy using 1-D optimization algorithms (e.g., dichotomous search method, golden section method, and Fibonacci search [32, pages 268-275]). The constraint that every ϵ_k lies in the interval [0, 1] can be easily enforced by using the interval [0, 1] as the initial interval of uncertainty for the chosen 1-D minimization method. Hence, the algorithm is guaranteed to provide detector efficiency estimates that lie in the interval [0, 1] when $\epsilon_k^0 \in [0, 1]$ and $\lambda_d^0 > 0$ for all k and d. We refer to the EM algorithm that results from using the CD method to solve $(\mathbf{P}_{\boldsymbol{\epsilon}})$ as the EM-CD algorithm. To solve the maximization problem in (2.18), we use the dichotomous search method [32].

EM-FP Algorithm

The second way we solve the problem $(\mathbf{P}_{\boldsymbol{\epsilon}})$ is to use a fixed-point iteration. Although it guarantees nonnegative detector efficiency estimates, the fixed-point iteration is not theoretically guaranteed to have the property that the estimates all lie in the interval [0, 1]. Nevertheless, in all of our simulations, the detector efficiency estimates obtained using the fixed-point iteration did have this property. An advantage of the fixed-point iteration is that it converges faster in practice than the coordinate descent algorithm.

Ignoring the constraint in $(\mathbf{P}_{\boldsymbol{\epsilon}})$, our approach is to use a fixed-point iteration to solve the system of equations

$$\frac{\partial g^{(m)}(\boldsymbol{\epsilon})}{\partial \epsilon_k} = 0, \quad k = 1, 2, \ldots, I. \tag{2.22}$$

Taking the partial derivative of $g^{(m)}$ with respect to ϵ_k gives

$$\frac{\partial g^{(m)}(\boldsymbol{\epsilon})}{\partial \epsilon_k} = \sum_{l \in \text{fan}_k} \left[\frac{b_{kl}}{\epsilon_k} + (n_{kl}^m - b_{kl}) \frac{-\epsilon_l}{1 - \epsilon_k \epsilon_l}\right]. \tag{2.23}$$

Thus, the resulting system of equations are

$$\sum_{l \in \text{fan}_k} [\frac{b_{kl}}{\epsilon_k} + (n_{kl}^m - b_{kl}) \frac{-\epsilon_l}{1 - \epsilon_k \epsilon_l}] = 0, \quad k = 1, 2, \ldots, I. \tag{2.24}$$

Obviously, a closed form solution for $\boldsymbol{\epsilon}^{m+1}$ is not possible. To find an approximate solution, we write (2.24) as

$$\sum_{l \in \text{fan}_k} b_{kl} (\frac{1}{\epsilon_k} + \frac{\epsilon_l}{1 - \epsilon_k \epsilon_l}) = \sum_{l \in \text{fan}_k} \frac{n_{kl}^m \epsilon_l}{1 - \epsilon_k \epsilon_l}, \quad k = 1, 2, \ldots, I. \tag{2.25}$$

Simplifying the left-hand side of (2.25) yields

$$\frac{1}{\epsilon_k} \sum_{l \in \text{fan}_k} \frac{b_{kl}}{1 - \epsilon_k \epsilon_l} = \sum_{l \in \text{fan}_k} \frac{n_{kl}^m \epsilon_l}{1 - \epsilon_k \epsilon_l}, \quad k = 1, 2, \ldots, I. \tag{2.26}$$

The following fixed-point iteration is then used to find an approximate solution to the system of equations in (2.26):

$$\epsilon_k^{m+1,i+1} = \frac{\sum_{l \in \text{fan}_k} \frac{b_{kl}}{1 - \epsilon_k^{m+1,i} \epsilon_l^{m+1,i}}}{\sum_{l \in \text{fan}_k} \frac{n_{kl}^m \epsilon_l^{m+1,i}}{1 - \epsilon_k^{m+1,i} \epsilon_l^{m+1,i}}}, \quad k = 1, 2, \ldots, I. \tag{2.27}$$

As with the EM-CD algorithm, the initial estimate of $\boldsymbol{\epsilon}^{m+1}$ is $\boldsymbol{\epsilon}^{m+1,0} \triangleq \boldsymbol{\epsilon}^m$. Further, the solution to $(\mathbf{P}_{\boldsymbol{\epsilon}})$ is the iterate obtained when the stopping criterion in (2.19) is satisfied.

If n_{kl}^m in (2.11) is positive for all (k, l), then every term in (2.27) is nonnegative. Thus, the detector efficiency estimates will be nonnegative provided $\mathbf{0} < \boldsymbol{\epsilon}^m \leq \mathbf{1}$ and $\boldsymbol{\lambda}^m > \mathbf{0}$ for all m. We refer to the EM algorithm that results from using the discussed fixed-point iteration as the EM-FP algorithm.

Over-Parameterized Approach

We now assume the unknowns are the set of parameters $\{\epsilon_{kl}\}$. This case is referred to as *over-parameterized* case because the number of unknowns in this case exceeds the number of data points.

Let $\boldsymbol{e}$ be the vector with elements $\{\epsilon_{kl}\}$. The conditional expectation of the log-likelihood function of the complete data given the incomplete data and the current estimates of $\boldsymbol{e}$ and $\boldsymbol{\lambda}$ is

$$\begin{aligned} U_o(\boldsymbol{e}, \boldsymbol{\lambda}; \boldsymbol{e}^m, \boldsymbol{\lambda}^m) &= \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [b_{kl} \log(\epsilon_{kl}) + (n_{kl}^m - b_{kl}) \log(1 - \epsilon_{kl})] \\ &+ \sum_{d=d_1}^{d_D} \sum_{\{(k,l):p(k,l)=d\}} [-\lambda_d + n_{kl}^m \log \lambda_d] + C_o, \end{aligned} \tag{2.28}$$

where C_o includes all the terms that are independent of $\boldsymbol{e}$ and $\boldsymbol{\lambda}$. The function $U_o(\boldsymbol{e}, \boldsymbol{\lambda}; \boldsymbol{e}^m, \boldsymbol{\lambda}^m)$ is decoupled:

$$U_o(\boldsymbol{e}, \boldsymbol{\lambda}; \boldsymbol{e}^m, \boldsymbol{\lambda}^m) = g_o^{(m)}(\boldsymbol{e}) + f_o^{(m)}(\boldsymbol{\lambda}) + C', \tag{2.29}$$

where

$$g_o^{(m)}(\boldsymbol{e}) = \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [b_{kl} \log(\epsilon_{kl}) + (n_{kl}^m - b_{kl}) \log(1 - \epsilon_{kl})] \tag{2.30}$$

and

$$f_o^{(m)}(\boldsymbol{\lambda}) = \sum_{d=d_1}^{d_D} \sum_{\{(k,l):p(k,l)=d\}} [-\lambda_d + n_{kl}^m \log \lambda_d]. \tag{2.31}$$

Because the function $U_o(\boldsymbol{e}, \boldsymbol{\lambda}; \boldsymbol{e}^m, \boldsymbol{\lambda}^m)$ is decoupled to two functions $g_o^{(m)}(\boldsymbol{e})$ and $f_o^{(m)}(\boldsymbol{\lambda})$, the problem of maximizing $U_o(\boldsymbol{e}, \boldsymbol{\lambda}; \boldsymbol{e}^m, \boldsymbol{\lambda}^m)$ with respect to $\boldsymbol{e}$ and $\boldsymbol{\lambda}$ is decoupled.

The E step for the over-parameterized formulation is the same as the E step for the sufficiently-parameterized formulation except that $\epsilon_k \epsilon_l$ is replaced by ϵ_{kl}. In other words, for $k = 1, 2, \ldots, I$, $l \in \text{fan}_k$, the conditional expectation of N_{kl} given the

incomplete data and the estimates $\boldsymbol{e}^m$ and $\boldsymbol{\lambda}^m$ is

$$n_{kl} \triangleq E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{e}^m, \boldsymbol{\epsilon}^m] = (1 - \epsilon_{kl}^m)\lambda_{p(k,l)}^m + b_{kl}. \tag{2.32}$$

However, the M step differs. New estimates for $\boldsymbol{e}$ and $\boldsymbol{\lambda}$ are obtained by maximizing $U_o(\boldsymbol{e}, \boldsymbol{\lambda}; \boldsymbol{e}^m, \boldsymbol{\lambda}^m)$ with respect to $\boldsymbol{e}$ and $\boldsymbol{\lambda}$ under the constraints

$$0 < \epsilon_{kl} \leq 1, \quad k = 1, 2, \ldots, I, \ l \in \text{fan}_k, \tag{2.33}$$

and

$$\lambda_d > 0, \quad d = d_1, d_2 \ldots, d_D. \tag{2.34}$$

Ignoring the constraints (2.33) and (2.34) momentarily, we take the derivatives of $g_o^{(m)}(\boldsymbol{e})$ and $f_o^{(m)}(\boldsymbol{\lambda})$ with respect to every ϵ_{kl} and λ_d, respectively, and set them to zero.

$$\frac{\partial g_o^{(m)}(\boldsymbol{e})}{\partial \epsilon_{kl}} = b_{kl}\frac{1}{\epsilon_{kl}} + (n_{kl}^m - b_{kl})\frac{-1}{1 - \epsilon_{kl}} = 0, \quad k = 1, 2, \ldots, I, \ l \in \text{fan}_k. \tag{2.35}$$

$$\frac{\partial f_o^{(m)}(\boldsymbol{\lambda})}{\partial \lambda_d} = \sum_{\{(k,l):p(k,l)=d\}} (-1 + n_{kl}^m \frac{1}{\lambda_d}) = 0, \quad d = d_1, d_2, \ldots, d_D. \tag{2.36}$$

Solving (2.35) and (2.36), we have the following iteration.

$$\epsilon_{kl}^{m+1} = \frac{b_{kl}}{n_{kl}^m}, \quad k = 1, 2, \ldots, I, \ l \in \text{fan}_k. \tag{2.37}$$

$$\lambda_d^{m+1} = \frac{\sum_{\{(k,l):p(k,l)=d\}} n_{kl}^m}{\text{number of pairs } (k,l) \text{ such that } p(k,l) = d}, \quad d = d_1, d_2, \ldots, d_D. \tag{2.38}$$

For the constraints (2.33) and (2.34) to be satisfied, ϵ_{kl}^0 must lie in the interval [0, 1] for $k = 1, 2, \ldots, I$, $l \in \text{fan}_k$ and λ_d^0 must be nonnegative for $d = d_1, d_2, \ldots, d_D$. As mentioned earlier, the results using this approach are poor. Consequently, we do not include this approach in our simulation studies.

2.2 Poisson Model II

As we mentioned earlier, the CRB for the estimators of $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$ in (**P1**) cannot be determined because the Fisher information matrix for the PDF in (**P1**) is non-invertible. One of the possible approaches to changing the non-invertibility is to reduce the number of unknowns in the estimation problem [35]. In this section, we present a second Poisson model that incorporates *a priori* information about the blank scan data so that the number of unknowns in the estimation problem is reduced.

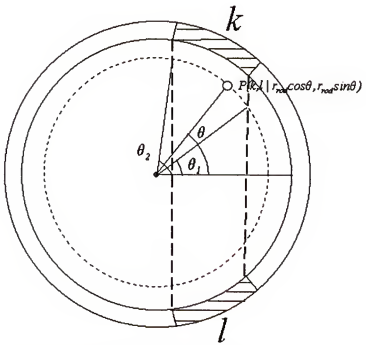


Figure 2.6: The exact way to calculate $\lambda_{p(k,l)}$, given $\lambda(r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta)$.

For detector pair (k,l) (see Fig. 2.6), the mean number of photon pairs that are transmitted along the detector pair, $\lambda_{p(k,l)}$, can be expressed as [37]

$$\lambda_{p(k,l)} = \int_{\text{rod path}} \lambda(x,y) P(k,l \mid x,y) dxdy, \tag{2.39}$$

where $\lambda(x,y)$ is the number of photon pairs emitted from the rod at point (x,y) and $P(k,l \mid x,y)$ is the probability that a photon pair emitted from the rod at point

(x, y) is incident on detector pair (k, l). Because the rod path is circular, it is more convenient to use polar coordinate than Cartesian coordinate. From the scanner geometry, $\lambda_{p(k,l)}$ can be expressed as

$$\lambda_{p(k,l)} = 2 \int_{\theta_1}^{\theta_2} \lambda(r_{\text{rod}} \cos\theta, r_{\text{rod}} \sin\theta) P(k, l | r_{\text{rod}} \cos\theta, r_{\text{rod}} \sin\theta) r_{\text{rod}} d\theta, \quad (2.40)$$

where $\lambda(r_{\text{rod}} \cos\theta, r_{\text{rod}} \sin\theta)$ is the mean number of photon pairs that are emitted from the rod at the angle θ, r_{rod} is the radius of the path of the rotating rod, and $P(k, l | r_{\text{rod}} \cos\theta, r_{\text{rod}} \sin\theta)$ is the probability that a photon pair emitted from the rod at the angle θ is incident upon the detector pair (k, l). The factor of two accounts for the fact that the rod passes through the same detector tube twice in one cycle. We assume that the rod rotates at a constant angular velocity and emits positron uniformly. Under these assumptions, it follows that $\lambda(r_{\text{rod}} \cos\theta, r_{\text{rod}} \sin\theta)$ is constant over θ. Let $\lambda \triangleq 2\lambda(r_{\text{rod}} \cos\theta, r_{\text{rod}} \sin\theta) r_{\text{rod}}$. Then, the integral in (2.40) may be approximated by the following sum (see Fig. 2.7):

$$\lambda_{p(k,l)} = \sum_n \lambda P(n\triangle\theta)\triangle\theta = \lambda c(k, l), \quad (2.41)$$

where

$$c(k, l) = \sum_n P_n \triangle\theta, \quad (2.42)$$

and $P_n \triangleq P(n\triangle\theta)$ is the probability that a photon pair emitted from the rod at the n^{th} position is incident upon the detector pair (k, l). This probability may be calculated using the angle-of-vies (AOV) method [14,36]. Specifically, the probability that a photon pair emitted from a point (x, y) is incident on detector pair (k, l) is determined by the angles α_1, α_2, α_3, and α_4 (see Fig. 2.8). The AOV is the maximum

angle for which a photon pair would be incident on a detector pair and is given by

$$\text{AOV} \triangleq \min\{\alpha_1, \alpha_2, \pi - \alpha_3, \pi - \alpha_4\}. \tag{2.43}$$

Then, the probability that a photon pair emitted from the point (x, y) is incident on detector pair (k, l) is equal to AOV/π [36]. Note, the parameter $c(k, l)$ depends

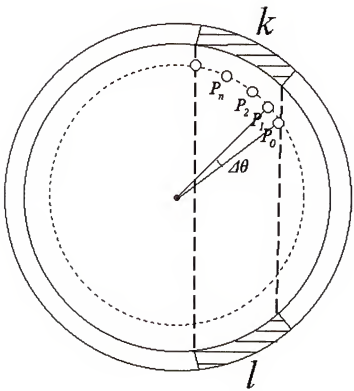


Figure 2.7: A proposed way to calculate $c(k, l)$ to approximate $\lambda_{p(k,l)}$.

only on $p(k, l)$. The parameters $\{c(k, l)\}$ for the Siemens-CTI ECAT EXACT 921 scanner are plotted in Fig. 2.9 as a function of $(l - k)$.

Equipped with the above information, we assume that N_{kl} and B_{kl} are Poisson with means $\lambda c(k, l)$ and $\epsilon_{kl}\lambda c(k, l)$, respectively:

$$N_{kl} \sim \text{Poisson}(\lambda c(k, l)),\ k = 1, 2, \ldots, I,\ l \in \text{fan}_k, \tag{2.44}$$

$$B_{kl} \sim \text{Poisson}(\epsilon_{kl}\lambda c(k, l)),\ k = 1, 2, \ldots, I,\ l \in \text{fan}_k. \tag{2.45}$$

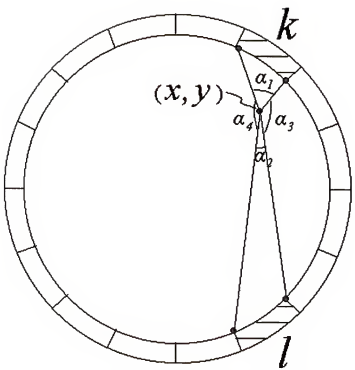


Figure 2.8: The four angles to determine the angle-of-view associated with the point (x, y) and detector pair (k, l).

We refer to the data model described in equations (2.44) and (2.45) as Poisson Model II.

2.2.1 Estimation Problem Associated with Poisson Model II

As discussed in the previous section, the parameters $\{c(k,l)\}$ are determined from the geometry of the scanner. The unknowns $\lambda_{d_1}, \lambda_{d_2}, \ldots, \lambda_{d_D}$ for Poisson Model I have been reduced to only one unknown, λ, for Poisson Model II.

The estimation problem now becomes

$$\textbf{(P2)} \qquad (\hat{\boldsymbol{\epsilon}}, \hat{\lambda}) = \arg\max_{\boldsymbol{\epsilon},\lambda} \phi_2(\boldsymbol{\epsilon}, \lambda) \text{ subject to } \mathbf{0} < \boldsymbol{\epsilon} \leq \mathbf{1} \text{ and } \lambda > 0, \qquad (2.46)$$

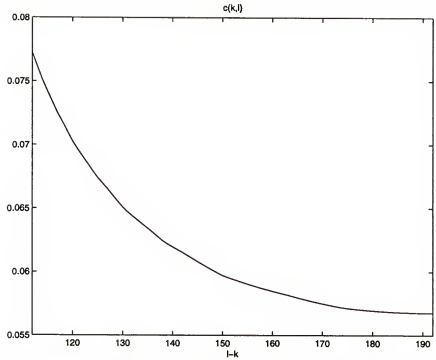


Figure 2.9: The parameters $\{c(k,l)\}$ for the Siemens-CTI ECAT EXACT 921 scanner calculated according to (2.42).

where $\phi_2(\boldsymbol{\epsilon}, \lambda)$ is the log-likelihood function:

$$\phi_2(\boldsymbol{\epsilon}, \lambda) \triangleq \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}, \lambda) = \log \prod_{k=1}^{I} \prod_{l \in \text{fan}_k} e^{-\epsilon_{kl}\lambda c(k,l)} \frac{(\epsilon_{kl}\lambda c(k,l))^{b_{kl}}}{b_{kl}!}. \tag{2.47}$$

We show in Appendix A that the function $\phi_2(\boldsymbol{\epsilon}, \lambda)$ is not concave.

Unfortunately, although the number of unknowns is reduced in **(P2)**, the Fisher information matrix for the probability density function associated with **(P2)** is still non-invertible (see Appendix B).

2.2.2 Estimation Algorithms Associated with Poisson Model II

We show in Appendix C that the conditional expectation of the log-likelihood function of the complete data $\boldsymbol{X} = [\boldsymbol{B}, \boldsymbol{N}]$ given the incomplete data and the current

estimates $\boldsymbol{\epsilon}^m$ and λ^m is

$$\begin{aligned} U'(\boldsymbol{\epsilon}, \lambda; \boldsymbol{\epsilon}^m, \lambda^m) &\triangleq E[L_c'(\boldsymbol{B}, \boldsymbol{N}; \boldsymbol{\epsilon}, \lambda) \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m] \\ &= \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [b_{kl} \log(\epsilon_k \epsilon_l) + (n_{kl}^m - b_{kl}) \log(1 - \epsilon_k \epsilon_l)] \\ &+ \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [-\lambda c(k,l) + n_{kl}^m \log \lambda c(k,l)] + C_2, \end{aligned} \tag{2.48}$$

where $L_c'(\boldsymbol{B}, \boldsymbol{N}; \boldsymbol{\epsilon}, \lambda)$ is the log-likelihood function of the complete data:

$$L_c'(\boldsymbol{B}, \boldsymbol{N}; \boldsymbol{\epsilon}, \lambda) \triangleq P(\boldsymbol{B} = \boldsymbol{b}, \boldsymbol{N} = \boldsymbol{n}; \boldsymbol{\epsilon}, \lambda), \tag{2.49}$$

and C_2 is independent of $\boldsymbol{\epsilon}$ and λ. We also show in Appendix C that, for $k = 1, 2, \ldots, I, l \in \text{fan}_k$,

$$n_{kl}^m \triangleq E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}^m, \lambda^m] = (1 - \epsilon_k^m \epsilon_l^m) \lambda^m c(k,l) + b_{kl}. \tag{2.50}$$

The function $U'(\boldsymbol{\epsilon}, \lambda; \boldsymbol{\epsilon}^m, \lambda^m)$ is given by

$$U'(\boldsymbol{\epsilon}, \lambda; \boldsymbol{\epsilon}^m, \lambda^m) = g^{(m)}(\boldsymbol{\epsilon}) + f_2^{(m)}(\lambda) + C_2, \tag{2.51}$$

where

$$g^{(m)}(\boldsymbol{\epsilon}) = \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [b_{kl} \log(\epsilon_k \epsilon_l) + (n_{kl}^m - b_{kl}) \log(1 - \epsilon_k \epsilon_l)] \tag{2.52}$$

and

$$f_2^{(m)}(\lambda) \triangleq \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [-\lambda c(k,l) + n_{kl}^m \log \lambda c(k,l)]. \tag{2.53}$$

The steps of the EM algorithm are:

Step I. (*Initialization*)

Set m=0.

Choose $\boldsymbol{\epsilon}^0$ *and* λ^0 *s.t.* $\mathbf{0} < \boldsymbol{\epsilon}^0 \leq \mathbf{1}$ *and* $\lambda^0 > 0$.

Step II. (*E step*)

$$n_{kl}^m = (1 - \epsilon_k^m \epsilon_l^m)\lambda^m c(k,l) + b_{kl}, \quad k = 1, 2, \ldots, I, \; l \in fan_k.$$

Step III. (*M step*)

$(\mathbf{P}_\lambda)$ $\lambda^{m+1} \triangleq \arg\max_{\lambda > 0} f_2^{(m)}(\lambda)$.

$(\mathbf{P}_{\boldsymbol{\epsilon}})$ $\boldsymbol{\epsilon}^{m+1} \triangleq \arg\max_{\mathbf{0} < \boldsymbol{\epsilon} \leq \mathbf{1}} g^{(m)}(\boldsymbol{\epsilon})$.

Step IV.

If stopping criterion is not satisfied, set $m = m + 1$ *and go to Step II.*

Otherwise stop.

The problem $(\mathbf{P}_{\boldsymbol{\epsilon}})$ is exactly the same optimization problem that arose when Poisson Model I was used. Therefore, the detector efficiencies can be estimated using (2.18) and (2.27). To find the solution to $(\mathbf{P}_\lambda)$, we ignore the constraint $\lambda > 0$ momentarily. Taking the derivative of $f_2^{(m)}(\lambda)$ with respect to λ and setting it to zero, we get

$$\sum_{k=1}^{I} \sum_{l \in \text{fan } k} [-c(k,l) + \frac{n_{kl}^m}{\lambda}] = 0. \tag{2.54}$$

Solving (2.54), the new estimate for λ is

$$\lambda^{m+1} = \frac{\sum_{k=1}^{I} \sum_{l \in \text{fan } k} n_{kl}^m}{\sum_{k=1}^{I} \sum_{l \in \text{fan } k} c(k,l)}. \tag{2.55}$$

We refer to the estimation algorithm using (2.18) and (2.55) for Poisson Model II as the EM-CD2 algorithm. We refer to the estimation algorithm using (2.27) and (2.55) for Poisson Model II as the EM-FP2 algorithm.

CHAPTER 3
ML ESTIMATION ALGORITHMS USING A REFINED DATA MODEL

3.1 Poisson Model III

To get a more complete data model for the blank scan data, the effect of detector penetration must be incorporated. In Fig. 3.1, although the photons are incident

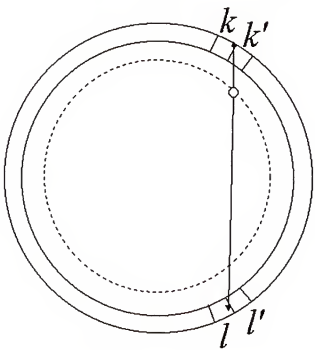


Figure 3.1: Detector penetration.

upon detector pair (k', l'), the photons may penetrate the detectors. Suppose the photon incident on detector k' penetrates detector k' and is detected by detector k, and the photon incident on detector l' penetrates detector l' and is detected by detector l. Then, a coincidence may be recorded erroneously by detector pair (k, l). It should be noted that a coincidence may also be erroneously recorded by detector

pairs (k, l') or (k', l) when detector penetration takes place, even though the photons are incident upon detector pair (k', l'). Even if the photons stop within detectors k and l, the photons might not be detected by detectors k and l. For a photon to be detected by a detector, the photon must have enough energy after stopping to interact with the detector [25].

The probability that the photon penetrates detector k' (or survival rate) is $e^{-L_{k'}\mu}$ [38, 39], where $L_{k'}$ is the length of intersection of detector k' and the path of the photon (see Fig. 3.2) and μ is the attenuation coefficient of the detector. Therefore,

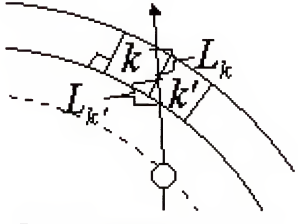


Figure 3.2: Definitions of L_k and $L_{k'}$.

for the given path of the photon pair, the probability that the photon incident on detector k' is detected by detector k is equal to

$$\rho_k e^{-L_{k'}\mu}(1 - e^{-L_k\mu}), \tag{3.1}$$

where ρ_k is the probability that a photon event is recorded given that a photon has been stopped by detector k, and $(1 - e^{-L_k\mu})$ is the probability that the photon is stopped by detector k. The attenuation coefficient of detectors is assumed to be known. For BGO detectors, the linear attenuation coefficient is 0.96 cm^{-1} [25] and the intersection lengths can be calculated from the scanner geometry. Although we

focused on one photon of the photon pair, similar statements can be made about the other photon.

When the effect of detector penetration is considered, the tube associated with a detector pair is defined differently (see Fig. 3.3). The mean number of photon pairs

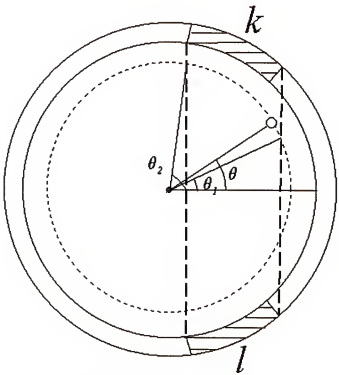


Figure 3.3: The tube associated with detectors k and l is defined by the dashed lines when detector penetration is taken into account. Note that, to clearly depict the angles, detectors k and l are magnified.

that are detected by detector pair (k,l) is given by

$$\begin{aligned} & E[B_{kl}] \\ &= \int_{\text{rod path}} \lambda(x,y)\rho_k\rho_l P(k,l|x,y)dxdy \\ &= 2\int_{\theta_1}^{\theta_2} \lambda(r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta)\rho_k\rho_l P(k,l|r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta)r_{\text{rod}}d\theta, \end{aligned} \tag{3.2}$$

where $\lambda(r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta)$ is the mean number of photon pairs that are emitted from the rod at angle θ, and $P(k,l|r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta)$ is the probability that a

photon pair emitted from the rod at angle θ is stopped within detector k and l, respectively. Since many paths are possible for a photon pair that is incident on detector pair (k,l), the probability $P(k,l|r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta)$ can be expressed as

$$\begin{aligned} &P(k,l|r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta) \\ &= \int_{\theta_3}^{\theta_3+\text{AOV}} e^{-L_{k'}(\alpha)\mu}(1-e^{-L_k(\alpha)\mu})e^{-L_{l'}(\pi+\alpha)\mu}(1-e^{-L_l(\pi+\alpha)\mu})\frac{d\alpha}{\pi}, \end{aligned} \tag{3.3}$$

where $L_k(\alpha)$ is the intersection length of the path of the photon traveling at angle α and detector k, and $L_{k'}(\alpha)$ is the intersection length of the path of the photon traveling at angle α and detector k' (see Figs. 3.4 and 3.5). The intersection lengths $L_l(\pi+\alpha)$ and $L_{l'}(\pi+\alpha)$ are defined similarly. Note that a photon might have

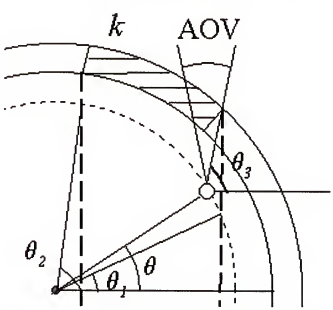


Figure 3.4: Angles that are used to calculated $P(k,l|r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta)$ in (3.3).

penetrated several detectors before reaching detector k. In this case, (3.3) would have to be modified accordingly.

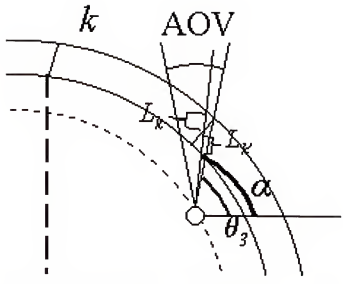


Figure 3.5: Angles that are used to calculated $P(k,l|r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta)$ in (3.3).

For convenience, we let $\lambda \triangleq 2\lambda(r_{\text{rod}}\cos\theta, r_{\text{rod}}\sin\theta)r_{\text{rod}}$. In practice, the exact integral in (3.2) cannot be determined. Therefore, we approximate the integral by

$$\lambda\rho_k\rho_l \sum_{n=1}^{N_\theta} P(n\triangle\theta)\triangle\theta, \tag{3.4}$$

where

$$P(n\triangle\theta) = \sum_{i=1}^{N_\alpha} e^{-L_{k'}(i\triangle\alpha)}(1-e^{-L_k(i\triangle\alpha)})e^{-L_{l'}(\pi+i\triangle\alpha)}(1-e^{-L_l(\pi+i\triangle\alpha)})\frac{\triangle\alpha}{\pi}, \tag{3.5}$$

$$\triangle\theta \triangleq (\theta_2-\theta_1)/N_\theta, \tag{3.6}$$

and

$$\triangle\alpha \triangleq \text{AOV}/N_\alpha. \tag{3.7}$$

The quantity in (3.4) can be expressed as

$$\lambda\rho_k\rho_l a(k,l), \tag{3.8}$$

where

$$a(k,l) = \sum_{n=1}^{N_\theta} P(n\triangle\theta)\triangle\theta. \tag{3.9}$$

Observe that $a(k,l)$ depends only on $p(k,l)$. The parameters $\{a(k,l)\}$ for the Siemens-CTI ECAT EXACT 921 scanner are plotted in Fig. 3.6. The number of photon pairs

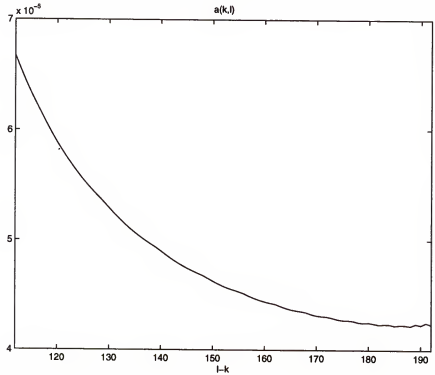


Figure 3.6: The parameters $\{a(k,l)\}$ for the Siemens-CTI ECAT EXACT 921 scanner calculated according to (3.9).

that are detected by detector pair (k,l) is therefore assumed to be an observation of the random variable B_{kl}, where

$$B_{kl} \sim \text{Poisson}(\lambda\rho_k\rho_l a(k,l)),\ k = 1,2,\ldots,I,\ l \in \text{fan}_k. \tag{3.10}$$

In order to utilize the same methodology for estimating unknown parameters of Poisson Model II, the means of random variables $\{B_{kl}\}$ must be related to the means of some other Poisson random variables in the same manners that the means of $\{B_{kl}\}$ are related to the means of $\{N_{kl}\}$. Let $\tilde{N}_{kl}$ be the number of photon pairs that are

stopped within detectors k and l. Then, the random variables $\{\tilde{N}_{kl}\}$ have means $\{\lambda a(k,l)\}$:

$$\tilde{N}_{kl} \sim \text{Poisson}(\lambda a(k,l)),\ k = 1, 2, \ldots, I,\ l \in \text{fan}_k. \tag{3.11}$$

We refer to the data model described in equations (3.10) and (3.11) as Poisson Model III.

3.1.1 Estimation Problem Associated with Poisson Model III

Since the parameters $\{a(k,l)\}$ can be determined from the scanner geometry (see (3.5) and (3.9)), the estimation problem now becomes

$$\textbf{(P3)} \qquad (\hat{\boldsymbol{\rho}}, \hat{\lambda}) = \arg\max_{\boldsymbol{\rho},\lambda} \phi_3(\boldsymbol{\rho}, \lambda) \text{ subject to } \mathbf{0} < \boldsymbol{\rho} \leq \mathbf{1} \text{ and } \lambda > 0, \tag{3.12}$$

where $\boldsymbol{\rho}$ is the vector with elements $\{\rho_k\}$ and $\phi_3(\boldsymbol{\rho}, \lambda)$ is the log-likelihood function:

$$\phi_3(\boldsymbol{\rho}, \lambda) \triangleq \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\rho}, \lambda) = \log \prod_{k=1}^{I} \prod_{l \in \text{fan}_k} e^{-\rho_k \rho_l \lambda a(k,l)} \frac{(\rho_k \rho_l \lambda a(k,l))^{b_{kl}}}{b_{kl}!}, \tag{3.13}$$

We show in Appendix A that the function $\phi_3(\boldsymbol{\rho}, \lambda)$ is not concave. In addition, we show in Appendix B that the Fisher information matrix for the probability density function associated with **(P3)** is non-invertible. Therefore, the Cramer-Rao lower bound for the estimators of $\boldsymbol{\rho}$ and λ cannot be determined.

3.1.2 Estimation Algorithms Associated with Poisson Model III

Because **Poisson Model III** has the exact same functional form as Poisson Model II, the derivation of estimation algorithms parallels that for **Poisson Model II**.

We show in Appendix C that the conditional expectation of the log-likelihood function of the complete data $\boldsymbol{X} = [\boldsymbol{B}, \boldsymbol{N}]$ given the incomplete data and the current

estimates $\boldsymbol{\rho}^m$ and λ^m is

$$\begin{aligned} U''(\boldsymbol{\rho}, \lambda; \boldsymbol{\rho}^m, \lambda^m) &\triangleq E[L_c''(\boldsymbol{B}, \boldsymbol{N}; \boldsymbol{\rho}, \lambda) \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\rho}^m, \lambda^m] \\ &= \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [b_{kl} \log(\rho_k \rho_l) + (n_{kl}^m - b_{kl}) \log(1 - \rho_k \rho_l)] \\ &+ \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [-\lambda a(k,l) + n_{kl}^m \log \lambda a(k,l)] + C_3, \end{aligned} \tag{3.14}$$

where $L_c''(\boldsymbol{B}, \boldsymbol{N}; \boldsymbol{\rho}, \lambda)$ is the log-likelihood function of the complete data:

$$L_c''(\boldsymbol{B}, \boldsymbol{N}; \boldsymbol{\rho}, \lambda) \triangleq \log P(\boldsymbol{B} = \boldsymbol{b}, \boldsymbol{N} = \boldsymbol{n}; \boldsymbol{\rho}, \lambda), \tag{3.15}$$

and where C_3 is independent of $\boldsymbol{\rho}$ and λ, We also show in Appendix C that, for $k = 1, 2, \ldots, I, l \in \text{fan}_k$,

$$n_{kl}^m \triangleq E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\rho}^m, \lambda^m] = (1 - \rho_k^m \rho_l^m) \lambda^m a(k,l) + b_{kl}, \tag{3.16}$$

The function $U''(\boldsymbol{\rho}, \lambda; \boldsymbol{\rho}^m, \lambda^m)$ is decoupled and given by

$$U''(\boldsymbol{\rho}, \lambda; \boldsymbol{\rho}^m, \lambda^m) = g^{(m)}(\boldsymbol{\rho}) + f_3^{(m)}(\lambda) + C_2, \tag{3.17}$$

where

$$g^{(m)}(\boldsymbol{\rho}) \triangleq \sum_{k=1}^{I} \sum_{l \in \text{fan } k} [b_{kl} \log(\rho_k \rho_l) + (n_{kl}^m - b_{kl}) \log(1 - \rho_k \rho_l)] \tag{3.18}$$

and

$$f_3^{(m)}(\lambda) \triangleq \sum_{d=d_1}^{d_D} \sum_{\{(k,l): p(k,l) = d\}} [-\lambda a(k,l) + n_{kl}^m \log \lambda a(k,l)]. \tag{3.19}$$

The steps of the EM algorithm are:

Step I. (*Initialization*)

Set m=0.

Choose $\boldsymbol{\rho}^0$ *and* λ^0 *s.t.* $\mathbf{0} < \boldsymbol{\rho}^0 \leq \mathbf{1}$ *and* $\lambda^0 > 0$.

Step II. (*E step*)

$n_{kl}^m = (1 - \rho_k^m \rho_l^m)\lambda^m a(k,l) + b_{kl}, \quad k = 1, 2, \ldots, I, \ l \in fan_k.$

Step III. (*M step*)

$(\mathbf{P}_\lambda)$ $\lambda^{m+1} \triangleq \arg\max_{\boldsymbol{\lambda}>0} f_3^{(m)}(\lambda).$

$(\mathbf{P}_{\boldsymbol{\rho}})$ $\boldsymbol{\rho}^{m+1} \triangleq \arg\max_{\mathbf{0}<\boldsymbol{\epsilon}\leq\mathbf{1}} g^{(m)}(\boldsymbol{\rho}).$

Step IV.

If stopping criterion is not satisfied, set $m = m + 1$ *and go to Step II.*

Otherwise stop.

Repeating the steps in Section 2.2.2, the problem $(\mathbf{P}_{\boldsymbol{\rho}})$ is solved either by using a CD-based method

$$\rho_k^{m+1,i+1} = \arg\max_{0<\rho\leq 1} f_2^{(m)}(\rho_1^{m+1,i+1}, \rho_2^{m+1,i+1}, \ldots, \rho_{k-1}^{m+1,i+1}, \rho, \rho_{k+1}^{m+1,i}, \ldots, \rho_I^{m+1,i}), \tag{3.20}$$

or a fixed-point iteration

$$\rho_k^{m+1,i+1} = \frac{\sum_{l \in \text{fan}_k} \frac{b_{kl}}{1-\rho_k^{m+1,i}\rho_l^{m+1,i}}}{\sum_{l \in \text{fan}_k} \frac{n_{kl}^m \rho_l^{m+1,i}}{1-\rho_k^{m+1,i}\rho_l^{m+1,i}}}, \tag{3.21}$$

where $\rho_k^{m+1,i+1}$ is the estimate of ρ_k^{m+1} at the $(i+1)^{th}$ iteration for $k = 1, 2, \ldots, I$. The new estimate for λ is given as

$$\lambda^{m+1} = \frac{\sum_{k=1}^{I} \sum_{l \in \text{fan}_k} n_{kl}^m}{\sum_{k=1}^{I} \sum_{l \in \text{fan}_k} a(k,l)}. \tag{3.22}$$

We refer to the estimation algorithm using (3.20) and (3.22) for **Poisson Model III** as the EM-CD3 algorithm. We refer to the estimation algorithm using (3.21) and (3.22) for Poisson Model III as the EM-FP3 algorithm.

CHAPTER 4
SIMULATION

In this chapter, we assess the performance of the proposed algorithms using synthetic PET data and data from a thorax phantom by Data Spectrum Inc. For the purpose of making comparisons, we apply the fan-sum algorithm and Ferreira's algorithm to this data as well. Note although Ferreira's algorithm was developed for a cylindrical source, the underlying assumptions could also be made for rotating rod sources.

For all the iterative algorithms, the initial estimates for the detector efficiencies are $\epsilon_k^0 = 0.5$ for all k. Given this choice, the initial estimate for $\boldsymbol{\lambda}$ for the EM-CD and EM-FP algorithms is

$$\lambda_d^0 = \frac{\sum_{\{(k,l):p(k,l)=d\}} \frac{b_{kl}}{c_k^0 c_l^0}}{\text{number of pairs } (k,l) \text{ s.t. } p(k,l) = d}, \; d = d_1, d_2, \ldots, d_D, \tag{4.1}$$

the initial estimate for λ for the EM-CD2 and EM-FP2 algorithms is

$$\lambda^0 = \frac{\sum_{k=1}^{I} \sum_{l \in \text{fan}_k} \frac{b_{kl}}{c_k^0 c_l^0}}{\sum_{k=1}^{I} \sum_{l \in \text{fan}_k} c(k,l)}, \tag{4.2}$$

and the initial estimate for λ for the EM-CD3 and EM-FP3 algorithms is

$$\lambda^0 = \frac{\sum_{k=1}^{I} \sum_{l \in \text{fan}_k} \frac{b_{kl}}{c_k^0 c_l^0}}{\sum_{k=1}^{I} \sum_{l \in \text{fan}_k} a(k,l)}. \tag{4.3}$$

The steps of the EM-CD and EM-FP algorithms are repeated until the following stopping criterion is met.

$$\sum_{k=1}^{I}(\frac{\epsilon_k^{m+1}-\epsilon_k^m}{\epsilon_k^m})^2+\sum_{d=d_1}^{d_D}(\frac{\lambda_d^{m+1}-\lambda_d^m}{\lambda_d^m})^2<\delta, \tag{4.4}$$

where δ is a small positive number. In our simulations, the EM-FP algorithm converged much faster than the EM-CD algorithm. Consequently, we used $\delta_\epsilon=\delta=10^{-11}$ for the EM-FP algorithm and $\delta_\epsilon=\delta=10^{-3}$ for the EM-CD algorithm. As for both the EM-CD2 and EM-FP2 algorithms, the stopping criterion is

$$\sum_{k=1}^{I}(\frac{\epsilon_k^{m+1}-\epsilon_k^m}{\epsilon_k^m})^2+(\frac{\lambda^{m+1}-\lambda^m}{\lambda^m})^2<\delta, \tag{4.5}$$

where δ is a small positive number. In our experiments, $\delta_\epsilon=\delta=10^{-3.5}$ for the EM-FP2 algorithm and $\delta_\epsilon=\delta=10^{-2.5}$ for the EM-CD2 algorithm. By contrast, Ferreira's algorithm did not converge even after 300,000 iterations when the value $\delta_\epsilon=10^{-1}$ was used. Therefore, in our simulations Ferreira's algorithm was executed for 250 iterations as specified by Ferreira et al. [17].

4.1 Synthetic Data

Using Poisson Model II, synthetic data are created to simulate blank scan data obtained from the Siemens-CTI ECAT EXACT 921 scanner. The scanner has 384 detectors per ring, and the ring diameter is 84.5 cm. The region to be reconstructed is a 43.9×43.9 cm^2 square consisting of $(128)^2$ voxels. Each plane consists of 192 projections and 160 members per projection. More information of this scanner can be found in Wienhard et al. [26]. As discussed in Chapter 2, this scanner has 81 possible perpendicular distances. The synthetic data consist of observations of Poisson random

variables $\{B_{kl}\}$:

$$B_{kl} \sim \text{Poisson}(\epsilon_k \epsilon_l \lambda c(k,l)), \quad k = 1, 2, \ldots, 384, l \in \text{fan}_k, \tag{4.6}$$

where

$$\lambda = 69000, \tag{4.7}$$

and $\{c(k,l)\}$ as discussed in Chapter 2. The value of λ in (4.7) is chosen so that the synthetic blank scan data corresponds to 2-hour scan. Uniform, piece-wise uniform, and nonuniform detector efficiencies are used in our experiments presented.

As customary [17], the detector efficiency estimates are normalized so that for each estimation algorithm the mean of the detector efficiency estimates is equal to one.

Because the EM-CD3 and EM-FP3 algorithms are based on Poisson Model III, which incorporates the effect of detector penetration, the EM-CD3 and EM-FP3 algorithms are not included. The reason is because synthetic data do not account for the effect of detector penetration. On the other hand, if the synthetic data did incorporate the effect of detector penetration, it would be unfair to compare the EM-CD3 and EM-FP3 algorithms against estimation algorithms that do not take detector penetration into consideration.

4.1.1 Comparison of Detector Efficiency Estimates

For our first experiment, the detector efficiencies were uniform:

$$\epsilon_k = 0.8, \quad k = 1, 2, \ldots, 384. \tag{4.8}$$

The true efficiencies are shown in Fig. 4.1(a), and the estimates provided by the fan-sum, Ferreira's, EM-CD, EM-FP, EM-CD2, and EM-FP2 algorithms are plotted

in Figs. 4.1(b)-4.1(g), respectively. From the estimation results, we conclude for all four of the proposed algorithms that the detector efficiency estimates have smaller variance. Note that the fan-sum estimates vary over a relatively wider range and the estimates provided by Ferreira's algorithm show a pronounced sinusoidal pattern.

For our second experiment, the detector efficiencies were piece-wise uniform:

$$\epsilon_k = \begin{cases} 0.8, & k = 1, 2, \ldots, 192 \\ 0.4, & k = 193, 194, 392 \end{cases} . \tag{4.9}$$

The true efficiencies are shown in Fig. 4.2(a), and the estimates provided by the fan-sum, Ferreira's, EM-CD, EM-FP, EM-CD2, and EM-FP2 algorithms are plotted in Figs. 4.2(b)-4.2(g), respectively. The fan-sum algorithm provides very poor estimates, while the estimates provided by Ferreira's algorithm again exhibits a sinusoidal pattern. From the experiments, the following conclusions are drawn about the proposed algorithms. The efficiency estimates for the first half of the detectors show little deviation from a constant value. This is also true for the second half of detectors. In addition, the ratio of the efficiency estimates for the first half of detectors to the efficiency estimates for the second half of detectors is very close to the true ratio 0.8/0.4.

To assess performance of estimation algorithms, we compare the mean-square error (MSE) of the efficiency estimates of 50 Monte Carlo (MC) runs. The error criterion MSE is defined as

$$\text{MSE} = \sum_{k=1}^{I} \frac{1}{50} \sum_{i=1}^{50} (\hat{\epsilon}_k(i) - \epsilon_k)^2, \tag{4.10}$$

where $\hat{\epsilon}_k(i)$ is the estimate for ϵ_k in the i^{th} MC run. The results are tabulated in Tables 4.1 and 4.2, respectively. Although each of four of the proposed algorithms provides detector efficiency estimates with smaller MSE than the fan-sum and Fer-

reira's algorithms, the difference between all algorithms is small. Further comparison is made based on a different criterion in the following.

Table 4.1: MSE of the detector efficiency estimates using synthetic data and uniform detector efficiencies in the first experiment.

Algorithm	Fan-sum	Ferreira's	EM-CD	EM-FP	EM-CD2	EM-FP2
MSE	15.3721	15.4538	15.3623	15.3623	15.3630	15.3630

Table 4.2: MSE of the detector efficiency estimates using synthetic data and piecewise uniform detector efficiencies in the second experiment.

Algorithm	Fan-sum	Ferreira's	EM-CD	EM-FP	EM-CD2	EM-FP2
MSE	69.3189	69.5227	68.0414	68.2702	68.2334	68.1167

Ideally, the detector efficiency estimates would be such that the ratio $\hat{\epsilon}_k/\epsilon_k$ is approximately equal to a constant for all k. Therefore, as a merit of figure, we use the variance of the ratio of the estimates and true values:

$$\text{VR} \triangleq \text{ sample variance of } \{\hat{\epsilon}_1/\epsilon_1, \hat{\epsilon}_2/\epsilon_2, \ldots, \hat{\epsilon}_I/\epsilon_I\}. \tag{4.11}$$

The idea is that the smaller the VR the better. To further compare the algorithms, 50 MC runs were performed for each of the first and second experiments and the VR was computed for each run. The plots of the VR versus MC runs are shown in Figs. 4.3 and 4.4 for the first and second experiments, respectively. From the plots, it can be clearly seen that all four of the proposed algorithms have comparable performance and each of them have much smaller VR than the fan-sum algorithm and Ferreira's algorithm.

For our third experiment, we used nonuniform detector efficiencies that were generated as follows:

$$\epsilon_k = \begin{cases} \xi_k & \text{if } 0 < \xi_k \leq 1 \\ 1 & \text{if } \xi_k > 1 \end{cases}, \quad k = 1, 2, \ldots, 384, \tag{4.12}$$

where

$$\xi_k = 0.5 + \sqrt{0.008}\mathcal{N}(0,1), \quad k = 1, 2, \ldots, 384. \tag{4.13}$$

Obviously, only positive values are used for $\{\xi_k\}$. The parameters in (4.13) are chosen so that with high probability the detector efficiencies lie in the interval [0.3, 0.7], which is a typical range for detector efficiencies.

The ratio of detector efficiency estimates to true detector efficiencies is plotted in Figs. 4.5(a)-4.5(f), where the detector efficiency estimates are provided by the fan-sum, Ferreira's, EM-CD, EM-FP, EM-CD2, and EM-FP2 algorithms, respectively. Based on Fig. 4.5, we confer that all four of the proposed algorithms provide much better detector efficiency estimates than the fan-sum and Ferreira's algorithms. Tabulated in Table 4.3 is the MSE of the detector efficiency estimates generated by using each of the estimation algorithms. All four of the proposed algorithms provide detector efficiency estimates with smaller MSE than the fan-sum and Ferreira's algorithms. We also compute the VR in each of 50 MC runs. The results are plotted in Fig. 4.6. It can be seen that all four of the proposed algorithms have comparable performance, and each of them is clearly better than the fan-sum algorithm and Ferreira's algo-

Table 4.3: MSE of the detector efficiency estimates using synthetic data and nonuniform detector efficiencies in the third experiment.

Algorithm	Fan-sum	Ferreira's	EM-CD	EM-FP	EM-CD2	EM-FP2
MSE	100.8640	100.7207	100.2472	100.2704	100.2215	100.2070

rithm in the sense that they provide estimates with smaller VR in every MC run than the fan-sum and Ferreira's algorithms.

Another comparison criterion adopted by other researchers is the relative error [9,17]:

$$\frac{|\hat{\epsilon}_k - \epsilon_k|}{\epsilon_k}, \quad k = 1, \ldots, I. \tag{4.14}$$

The maximum relative errors are plotted in Fig. 4.7, and the average relative errors are plotted separately in Fig. 4.8. Based on the maximum relative error, all four of the proposed algorithms have comparable performance and each of them is better than the fan-sum and Ferreira's algorithms. In terms of mean relative error, the EM-FP algorithm has the best performance, and both the EM-CD and EM-CD2 algorithms are better than the fan-sum and Ferreira's algorithms in every MC run. To remove the bias factor from the comparisons, we normalize the detector efficiency estimates so that their mean is equal to the mean of the true detector efficiencies and then compute the relative errors. Note that in practice this cannot be done because the true detector efficiencies are not known. The maximum and average relative errors associated with the normalized detector efficiency estimates are plotted in Figs. 4.9 and 4.10. The plots confirm that all four of the proposed algorithms have comparable performance, and each of them outperform the fan-sum and Ferreira's algorithms. Specifically, the proposed algorithms provide estimates with smaller maximum and average absolute errors than the fan-sum algorithm and Ferreira's algorithm in every MC run.

4.1.2 Comparison of Reconstructed Emission Images

The efficiency estimates were then used to reconstruct a synthetic emission image. Assuming only errors due to detector inefficiency, the emission data are assumed to be a realization of the set of Poisson random variables $\{Y_{kl}\}$ given in (1.3). For

convenience, the data model for $\{Y_{kl}\}$ is expressed again. The emission data, $\{y_{kl}\}$, are modeled as observations of the independent random variables $\{Y_{kl}\}$:

$$Y_{kl} \sim \text{Poisson}(\epsilon_k \epsilon_l [\boldsymbol{P}\boldsymbol{x}]_{kl}), \quad k = 1, 2, \ldots, 384, l \in \text{fan}_k,$$

where

$$[\boldsymbol{P}\boldsymbol{x}]_{kl} \triangleq \sum_{j=1}^{J} P_{kl,j} x_j, \tag{4.15}$$

and x_j is the mean number of photon pairs emitted from voxel j and J is the number of voxels. In addition, $P_{kl,j}$ is the probability that a photon pair emitted from voxel j is incident on detector pair (k,l) and $\boldsymbol{x} \triangleq [x_1, x_2, \ldots, x_J]^T$ is referred to as the *emission intensity vector*. Given the observed emission data $\boldsymbol{y}$, which is a vector with elements $\{y_{kl}\}$, the problem is to estimate the emission intensity vector $\boldsymbol{x}$. Given the synthetic emission intensity $\boldsymbol{x}$ in Fig. 4.11, we use the EMML algorithm [1,14,40,41] to estimate $\boldsymbol{x}$ from $\boldsymbol{y}$:

$$\hat{x}_j^{(n+1)} = \frac{\hat{x}_j^{(n+1)}}{\sum_{k,l} P^c_{kl,j}} \sum_{k,l} \frac{P^c_{kl,j} y_{kl}}{[\boldsymbol{P}^c \hat{\boldsymbol{x}}^{(n)}]_{kl}}, \tag{4.16}$$

where

$$P^c_{kl,j} = \epsilon_k \epsilon_l P_{kl,j}. \tag{4.17}$$

The efficiency estimates were used in (4.17) to reconstruct the emission image.

In Figs. 4.12(a)-4.12(d) are the plots of the reconstructed emission images at the 15^{th} iteration of the EMML algorithm using true detector efficiencies, no normalization, fan-sum algorithm, and Ferreira's algorithm, respectively. (The EMML algorithm virtually converges after 15 iterations.) And in Figs. 4.13(a)-4.13(d) are the plots of the reconstructed emission images at the 15^{th} iteration of the EMML algorithm using the EM-CD, EM-FP, EM-CD2, and EM-FP2 algorithms, respectively. Because all the reconstructed images look very similar to each other, 50 MC runs were used to compare the MSE of the reconstructed images. The criterion MSE is

defined for each iteration of the EMML algorithm as

$$\text{MSE} = \sum_{j=1}^{J} \frac{1}{50} \sum_{i=1}^{50} (\hat{x}_j(i) - x_j)^2, \tag{4.18}$$

where $\hat{x}_j(i)$ is the estimate for x_j at the i^{th} MC run. The curves of the MSE of reconstructed images versus iteration number of the EMML algorithm are shown in Fig. 4.14. Because the mean of true detector efficiencies is different from the mean of estimates of each estimation algorithm, bias becomes an issue and the MSE corresponding to true detector efficiencies is not included in the plot. While all

Table 4.4: MSE of the reconstructed emission images from synthetic data.

Iteration of EMML	16	17	18	19	20
Fan-sum	786.1761	785.5547	788.4473	794.3706	802.9343
Ferreira's	786.1472	785.5140	788.3954	794.3079	802.8612
EM-CD	786.1328	785.4880	788.3587	794.2610	802.8044
EM-FP	786.1363	785.4823	788.3440	794.2375	802.7723
EM-CD2	786.2022	785.3626	788.0479	793.7720	802.1426
EM-FP2	786.1835	785.2883	787.9220	793.5967	801.9190

the estimation algorithms improve the reconstructed emission images over using no normalization, the curves for all the estimation algorithms are close together. The values of MSE are tabulated in Table 4.4 for iterations 16-20 of the EMML algorithm for clarity.

4.2 Real Data

Real PET data were obtained from Dr. John Votaw of the PET Laboratory at the Emory University School of Medicine in Atlanta, GA. The PET scanner is the Siemens-CTI ECAT EXACT 921 model. The collected data consist of 47 planes, where each plane consists of 192 projections and 160 members per projection. The

scanner has 384 detectors per ring and the region to be reconstructed is a 43.9×43.9 cm^2 square consisting of $(128)^2$ voxels.

The fan-sum method, Ferreira's method, and the proposed algorithms were used to estimate the detector efficiencies for plane 10. The duration of the blank scan was 30 minutes. Since the true detector efficiencies are unknown, we have no objective way to compare the detector efficiency estimates. The efficiency estimates were used in conjunction with the EMML algorithm to reconstruct the emission image for plane 10. The duration of the emission scan was 7 minutes.

In Figs. 4.15(a)-(d) and 4.16 (a)-(b), the reconstructed emission images at the 20^{th} iteration of the EMML algorithm using the fan-sum, Ferreira's, EM-CD, EM-FP, EM-CD2, EM-FP2 algorithms, respectively, are shown. The reconstructed emission images are indistinguishable. As the EM-CD3 and EM-FP3 algorithms are based on Poisson Model III which incorporates the effect of detector penetration, when applying the EM-CD3 and EM-FP3 algorithms to reconstruct emission images, the probability matrix $\boldsymbol{P}$ in (1.3) must be defined differently. Let $P_{kl,j}$ stand for the probability that a photon pair is stopped by detectors k and l given that the photon pair is emitted from voxel j. The methods by Shepp et al. [14] and Carroll [36] can be extended to compute these probabilities [42]. The reconstructed emission images using the EM-CD3, EM-FP3 algorithms and no normalization (i.e., $\{\hat{\rho}_k = 1\}$) are plotted in Figs 4.17(a)-(c). The reconstructed emission images for the EM-CD3 and EM-FP3 algorithms show details (e.g. heart, lungs, and spine) much more clearly than the reconstructed emission images for the other estimation algorithms. Nevertheless, it is difficult to objectively determine whether using the EM-CD3 and EM-FP3 algorithms lead to better reconstructed emission images than using no normalization.

To date, there is no accepted methodology for assessing the performance of detector efficiency estimation algorithms when real data are used. Dr. Votaw suggested the following experiment. A cylindrical phantom with uniform activity could be scanned

for an extremely long time ($>$ 48 hours). Then, the emission images corresponding to each detector efficiency estimation algorithm would be reconstructed. Finally, the average variance of the pixel values for each image would be computed. The detector efficiency estimation algorithm that gave rise to the smallest variance would be considered to be the best algorithm.

The shortcoming of the experiment described above is that only a single realization of the blank scan data is used. To account for the statistical properties of the detector efficiency estimation algorithms, multiple realizations of the blank scan data would be needed. Unfortunately, it is difficult to obtain multiple realizations because of the extremely high scanning costs.

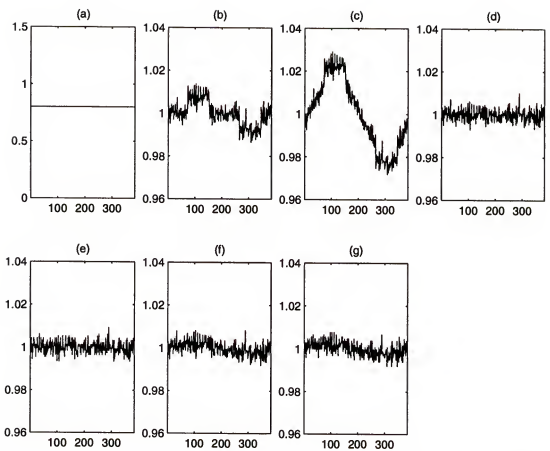


Figure 4.1: True efficiencies of synthetic data in the first experiment and the estimates: (a) true efficiencies; the others are the detector efficiency estimates generated by using (b) fan-sum algorithm; (c) Ferreira's algorithm; (d) EM-CD algorithm; (e) EM-FP algorithm; (f) EM-CD2 algorithm; and (g) EM-FP2 algorithm.

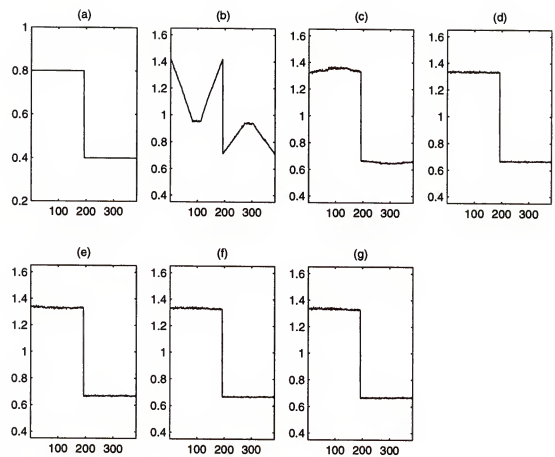


Figure 4.2: True efficiencies of synthetic data in the second experiment and the estimates: (a) true efficiencies; the others are the detector efficiency estimates generated by using (b) fan-sum algorithm; (c) Ferreira's algorithm; (d) EM-CD algorithm; (e) EM-FP algorithm; (f) EM-CD2 algorithm; and (g) EM-FP2 algorithm.

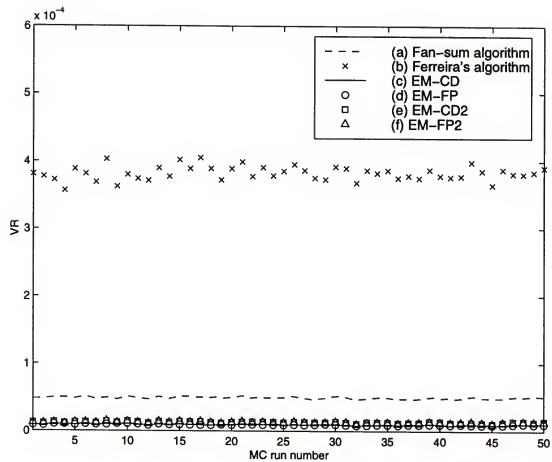


Figure 4.3: VR of the detector efficiency estimates in the first experiment versus MC run number using synthetic data, uniform detector efficiencies, and: (a) fan-sum algorithm; (b) Ferreira's algorithm; (c) EM-CD algorithm; (d) EM-FP algorithm; (e) EM-CD2 algorithm; and (f) EM-FP2 algorithm.

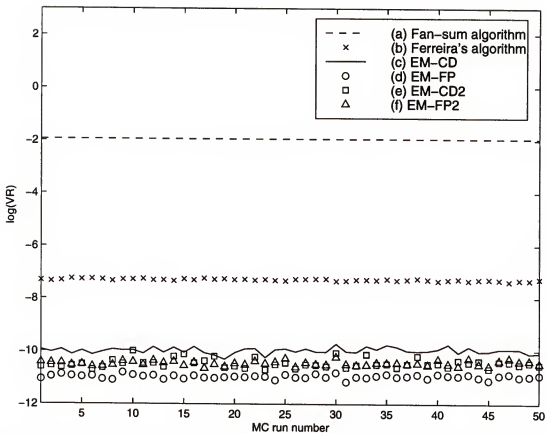


Figure 4.4: VR of the detector efficiency estimates in the second experiment versus MC run number using synthetic data, piece-wise uniform detector efficiencies, and: (a) fan-sum algorithm; (b) Ferreira's algorithm; (c) EM-CD algorithm; (d) EM-FP algorithm; (e) EM-CD2 algorithm; and (f) EM-FP2 algorithm.

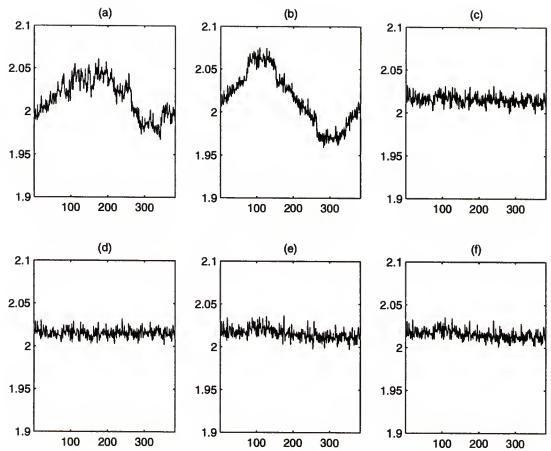


Figure 4.5: The ratio of the detector efficiency estimates and true detector efficiencies, where the detector efficiency estimates are generated using synthetic data, nonuniform detector efficiencies, and: (a) fan-sum algorithm; (b) Ferreira's algorithm; (c) EM-CD algorithm; (d) EM-FP algorithm; (e) EM-CD2 algorithm; and (f) EM-FP2 algorithm.

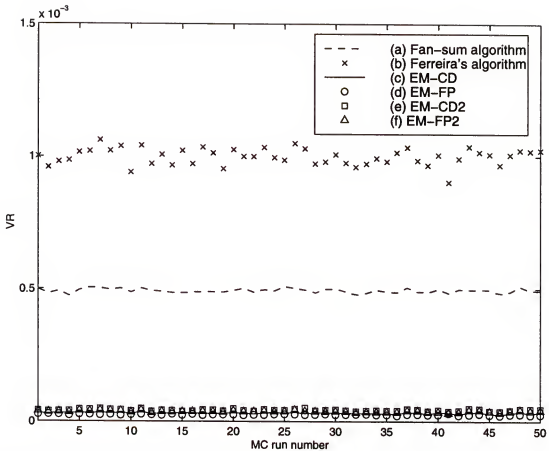


Figure 4.6: VR of the detector efficiency estimates in the third experiment versus MC run number using synthetic data, nonuniform detector efficiencies, and: (a) fan-sum algorithm; (b) Ferreira's algorithm; (c) EM-CD algorithm; (d) EM-FP algorithm; (e) EM-CD2 algorithm; and (f) EM-FP2 algorithm.

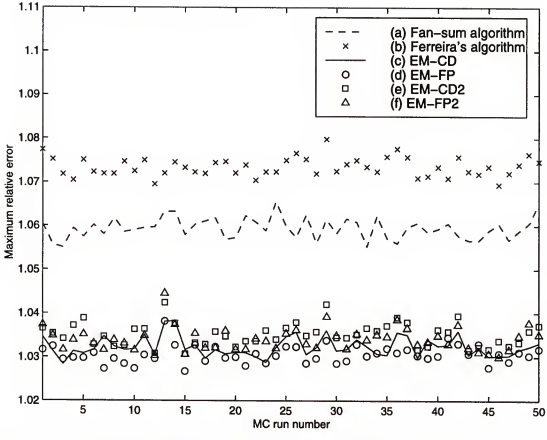


Figure 4.7: Maximum relative error of the detector efficiency estimates versus MC run number using synthetic data, nonuniform detector efficiencies, and: (a) fan-sum algorithm; (b) Ferreira's algorithm; (c) EM-CD algorithm; (d) EM-FP algorithm; (e) EM-CD2 algorithm; and (f) EM-FP2 algorithm.

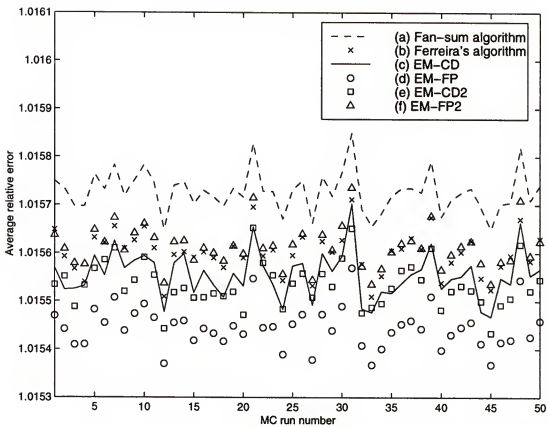


Figure 4.8: Average relative error of the detector efficiency estimates versus MC run number using synthetic data, nonuniform detector efficiencies, and: (a) fan-sum algorithm; (b) Ferreira's algorithm; (c) EM-CD algorithm; (d) EM-FP algorithm; (e) EM-CD2 algorithm; and (f) EM-FP2 algorithm.

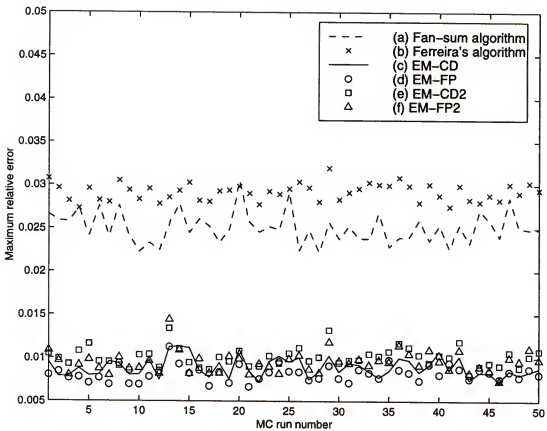


Figure 4.9: Maximum relative error of the detector efficiency estimates versus MC run number using synthetic data, nonuniform detector efficiencies, and: (a) fan-sum algorithm; (b) Ferreira's algorithm; (c) EM-CD algorithm; (d) EM-FP algorithm; (e) EM-CD2 algorithm; and (f) EM-FP2 algorithm. Note that the detector efficiency estimates are normalized so that their mean equals the mean of true efficiencies.

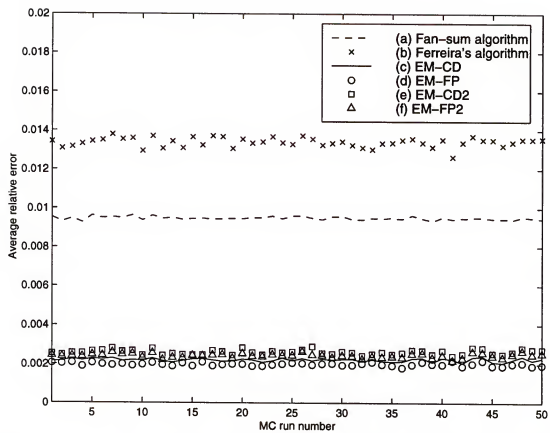


Figure 4.10: Average relative error of the detector efficiency estimates versus MC run number using synthetic data, nonuniform detector efficiencies, and: (a) fan-sum algorithm; (b) Ferreira's algorithm; (c) EM-CD algorithm; (d) EM-FP algorithm; (e) EM-CD2 algorithm; and (f) EM-FP2 algorithm. Note that the detector efficiency estimates are normalized so that their mean equals the mean of true efficiencies.

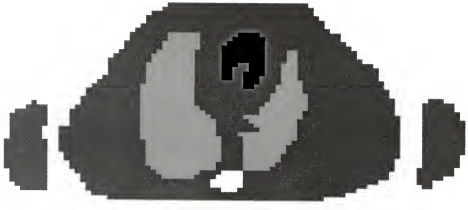


Figure 4.11: Synthetic thorax phantom.

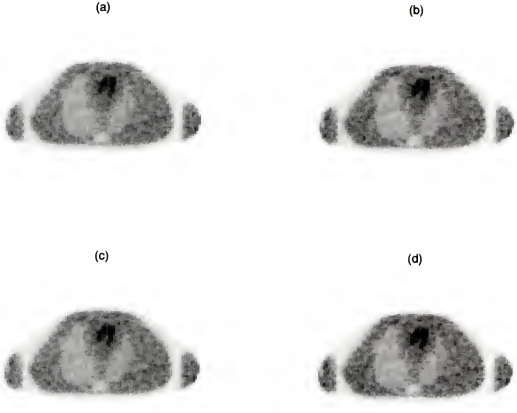


Figure 4.12: Reconstructed images of the thorax phantom in Fig. 4.11 at the 15^{th} iteration of the EMML algorithm using: (a) true efficiencies; (b) no normalization; (c) fan-sum algorithm; and (d) Ferreira's algorithm.

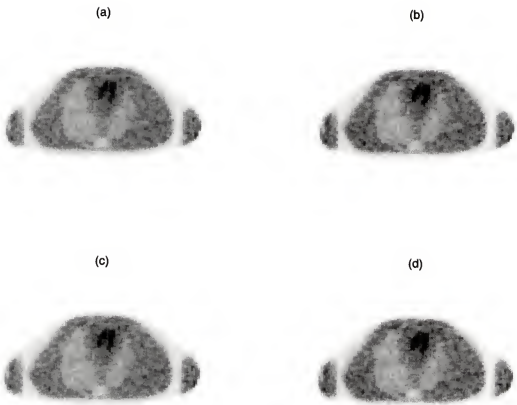


Figure 4.13: Reconstructed images of the thorax phantom in Fig. 4.11 at the 15^{th} iteration of the EMML algorithm using: (a) EM-CD algorithm; (b) EM-FP algorithm; (c) EM-CD2 algorithm; and (d) EM-FP2 algorithm.

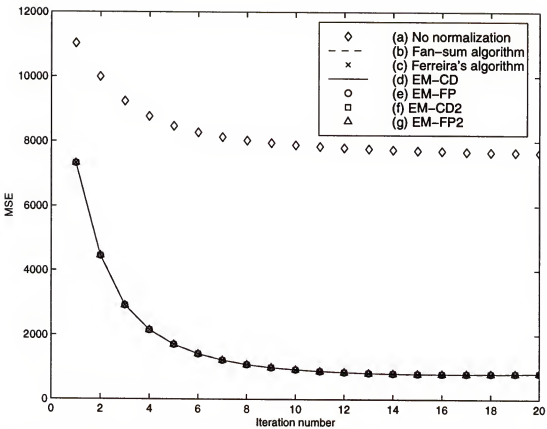


Figure 4.14: MSE of the reconstructed emission images versus the iteration number of the EMML algorithm using: (a) no normalization; (b) fan-sum algorithm; (c) Ferreira's algorithm; (d) EM-CD algorithm; (e) EM-FP algorithm; (f) EM-CD2 algorithm; and (g) EM-FP2 algorithm.

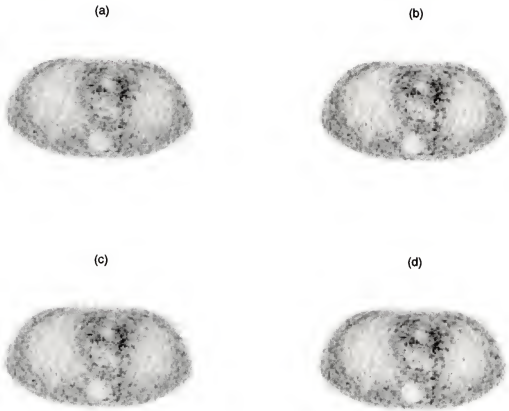


Figure 4.15: Reconstructed emission images of plane 10 of a thorax phantom at the 20^{th} iteration of EMML algorithm using: (a) the fan-sum algorithm; (b) Ferreira's algorithm; (c) EM-CD algorithm; and (d) EM-FP algorithm.

Figure 4.16: Reconstructed emission images of plane 10 of a thorax phantom at the 20^{th} iteration of EMML algorithm using: (a) EM-CD2 algorithm; and (b) EM-FP2 algorithm.

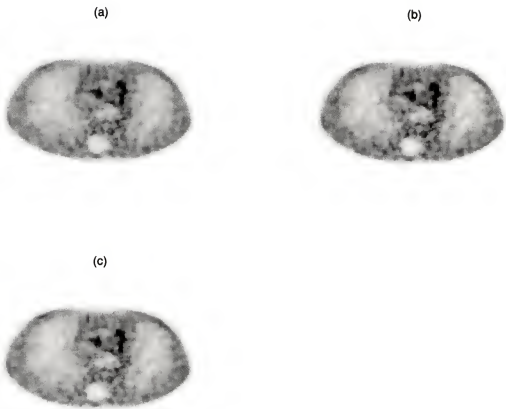


Figure 4.17: Reconstructed emission images of plane 10 of a thorax phantom at the 20^{th} iteration of EMML algorithm using: (a) EM-CD3 algorithm; (b) EM-FP3 algorithm; and (c) no normalization (i.e., $\{\hat{\rho}_k = 1\}$).

CHAPTER 5
CONCLUSIONS

Algorithms for computing ML estimates of detector efficiencies in PET have been proposed. The algorithms follow from the EM algorithm and three Poisson models that were developed for the PET blank scan data. The refined Poisson model (i.e., Poisson Model III) directly accounts for the Poisson nature of blank scan data, certain geometric properties of scanners, and detector penetration. The close resemblance between real data and the synthetic data generated according to the proposed models provides some justification for the models.

For each Poisson model, a closed form expression for the E step of the EM algorithm was derived. The corresponding M step was solved iteratively using the coordinate descent method and a fixed-point iteration. The detector efficiency estimation algorithms that were based on the coordinate descent method have the desirable property that the estimates were guaranteed to lie in the interval [0, 1]. The fixed-point iteration based algorithms do not have this property. However, they converge about 20 times faster than the coordinate descent based algorithms. All of the algorithms have the property that the corresponding log-likelihood function increase monotonically with increasing iterations.

Experiments using synthetic and real data were used to assess the proposed algorithms. For the synthetic data case, the following objective criteria were used: (1) variance of the ratio of the efficiency estimates and true efficiencies, (2) maximum relative error, and (3) average relative error. In terms of these criteria, the proposed algorithms greatly outperformed the fan-sum and Ferreira's algorithms. However, the reconstructed emission images for all of the estimation algorithms were visually iden-

tical. In addition, the mean-square error (MSE) values associated with the emission images were also similar, but the proposed algorithms did produce the smallest MSE.

There is no accepted approach for comparing the performance of detector efficiency estimation algorithms when real data are used. We expect the EM-CD3 and EM-FP3 algorithms to have a large impact because Poisson Model III takes the effect of detector penetration into consideration. The reconstructed emission images for the fan-sum, Ferreira's, EM-CD, EM-FP, EM-CD2, and EM-FP2 algorithms are indistinguishable. When applying the EM-CD3 and EM-FP3 algorithms to reconstruct emission images, a different probability matrix $\boldsymbol{P}$ is used. The reconstructed emission images for the EM-CD3 and EM-FP3 algorithms show details much more clearly than the reconstructed emission images for the other estimation algorithms. While it is difficult to objectively determine whether using the EM-CD3 and EM-FP3 algorithms lead to better reconstructed emission images than using no normalization, we conclude that the effect of detector penetration is a significant factor when reconstructing emission images.

CHAPTER 6
FUTURE WORK

While the proposed algorithms converge in practice, the convergence properties of the proposed algorithms are not well understood. Theoretic analysis of the proposed algorithm will be pursued to determine their convergence properties.

As we stated in previous chapters, the performance of the proposed algorithms depends on the initialization step because the objective functions are not concave. It is a common practice to use a constant value as the initial detector efficiency estimate for iterative algorithms (e.g., Ferreira's algorithm [17]). From our experiments, we found that the average values of the detector efficiency estimates obtained from the proposed algorithms are very close to the constant value that was used to initialize the algorithms. However, if the constant value is scaled, detector efficiency estimates are not correspondingly scaled. Therefore, the effect of the initialization step must be investigated.

APPENDIX A
CONCAVITY OF THE LOG-LIKELIHOOD FUNCTIONS

A.1 Poisson Model I

The log-likelihood function for the blank scan data is given by

$$\begin{aligned}\phi_1(\epsilon,\lambda) &\triangleq \log P(\boldsymbol{B}=\boldsymbol{b};\epsilon,\lambda)\\ &= \sum_k \sum_{l\in \text{fan}_k} [-\epsilon_k\epsilon_l\lambda_{p(k,l)} + b_{kl}\log(\epsilon_k\epsilon_l\lambda_{p(k,l)})]\\ &= \sum_d \sum_{\{(k,l):p(k,l)=d\}} [-\epsilon_k\epsilon_l\lambda_d + b_{kl}\log(\epsilon_k\epsilon_l\lambda_d)]. \end{aligned} \tag{A.1}$$

To determine whether a function is concave or not, we check if its Hessian is negative definite. To determine if a hermitian matrix is negative definite, we use the following theorem [43].

Theorem 1 *The following condition is necessary and sufficient for an $n\times n$ hermitian $\boldsymbol{H}$ to be negative definite. The principal minors consisting of the the determinants of the $k\times k$ matrices in the top left-hand corner of $\boldsymbol{H}$ $(k=1,\ldots,n)$ are all negative:*

$$H_{11}<0,\quad \begin{vmatrix} H_{11} & H_{12}\\ H_{21} & H_{22}\end{vmatrix} < 0,\quad \begin{vmatrix} H_{11} & H_{12} & H_{13}\\ H_{21} & H_{22} & H_{23}\\ H_{31} & H_{32} & H_{33}\end{vmatrix} < 0,\ldots$$

The (i,j) element of Hessian for the log-likelihood function is

$$H_{ij} = \frac{\partial^2\phi_1(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}, \tag{A.2}$$

where $\boldsymbol{\theta} = [\boldsymbol{\epsilon}, \boldsymbol{\lambda}]$. The first partial derivatives of $\phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda})$ are

$$\frac{\partial \phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda})}{\partial \epsilon_k} = \sum_{l \in \text{fan}_k} [-\epsilon_l \lambda_{p(k,l)} + \frac{b_{kl}}{\epsilon_k}], \; k = 1, 2, \ldots, I,$$

and

$$\frac{\partial \phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda})}{\partial \lambda_d} = \sum_{\{(k,l): p(k,l) = d\}} [-\epsilon_k \epsilon_l + \frac{b_{kl}}{\lambda_d}], \; d = d_1, d_2, \ldots, d_D.$$

And, the second partial derivatives are

$$\begin{aligned}
\frac{\partial^2 \phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda})}{\partial \epsilon_k \partial \epsilon_l} &= \begin{cases} -\sum_{l \in \text{fan}_k} \frac{b_{kl}}{\epsilon_k^2} & \text{if } k = l \\ -\lambda_{p(k,l)} & \text{if } l \in \text{fan}_k \\ 0 & \text{otherwise} \end{cases}, \\
\frac{\partial^2 \phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda})}{\partial \lambda_d \partial \epsilon_k} &= -\sum_{l \in \text{fan}_k} [\epsilon_l \delta(p(k,l) - d)], \\
\frac{\partial^2 \phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda})}{\partial \lambda_d \partial \lambda_m} &= \begin{cases} \sum_{\{(k,l): p(k,l) = d\}} [-\frac{b_{kl}}{\lambda_d^2}] & \text{if } d = m \\ 0 & \text{otherwise} \end{cases}.
\end{aligned} \tag{A.3}$$

From (A.3), we know that

$$H_{11} = -\frac{1}{\epsilon_1^2} \sum_{l \in \text{fan}_1} b_{1l} < 0,$$

$$H_{12} = H_{21} = 0,$$

and

$$H_{22} = -\frac{1}{\epsilon_2^2} \sum_{l \in \text{fan}_2} b_{2l} < 0.$$

Therefore,

$$\begin{vmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{vmatrix} = H_{11} H_{22} - H_{12} H_{21} > 0.$$

So we conclude that $\boldsymbol{H}$ is not negative definite and that the objective function of $(\mathbf{P1})$, $\phi_1(\boldsymbol{\epsilon}, \boldsymbol{\lambda})$, is not concave.

For the *over-parameterized* approach, the same conclusion can be reached by a similar derivation. The log-likelihood function for the blank scan data is given by

$$\begin{aligned}
\log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{e}, \boldsymbol{\lambda}) &= \log \prod_k \prod_{l \in \text{fan}_k} e^{-\epsilon_{kl}\lambda_{p(k,l)}} \frac{(\epsilon_{kl}\lambda_{p(k,l)})^{b_{kl}}}{b_{kl}!} \\
&= \log \prod_d \prod_{\{(k,l):p(k,l)=d\}} e^{-\epsilon_{kl}\lambda_d} \frac{(\epsilon_{kl}\lambda_d)^{b_{kl}}}{b_{kl}!} \\
&= \sum_k \sum_{l \in \text{fan}_k} [-\epsilon_{kl}\lambda_{p(k,l)} + b_{kl} \log(\epsilon_{kl}\lambda_{p(k,l)}) + \log(b_{kl}!)] \\
&= \sum_d \sum_{\{(k,l):p(k,l)=d\}} [-\epsilon_{kl}\lambda_d + b_{kl} \log(\epsilon_{kl}\lambda_d) + \log(b_{kl}!)]. \quad \text{(A.4)}
\end{aligned}$$

The first partial derivatives of log-likelihood function are

$$\frac{\partial \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{e}, \boldsymbol{\lambda})}{\partial \epsilon_{kl}} = -\lambda_{p(k,l)} + \frac{b_{kl}}{\epsilon_{kl}}, \; k = 1, 2, \ldots, I, \; l \in \text{fan}_k, \qquad \text{(A.5)}$$

and

$$\frac{\partial \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{e}, \boldsymbol{\lambda})}{\partial \lambda_d} = \sum_{\{(k,l):p(k,l)=d\}} [-\epsilon_{kl} + \frac{b_{kl}}{\lambda_d}], \; d = d_1, d_2, \ldots, d_D. \qquad \text{(A.6)}$$

Further, the second partial derivatives are

$$\frac{\partial^2 \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{e}, \boldsymbol{\lambda})}{\partial \epsilon_{kl} \partial \epsilon_{mn}} = \begin{cases} -\frac{b_{kl}}{\epsilon_{kl}^2} & \text{if } (k,l) = (m,n) \\ 0 & \text{if } (k,l) \neq (m,n) \end{cases}, \qquad \text{(A.7)}$$

$$\frac{\partial^2 \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{e}, \boldsymbol{\lambda})}{\partial \lambda_d \partial \epsilon_{kl}} = \begin{cases} -1 & \text{if } p(k,l) = d \\ 0 & \text{if } p(k,l) \neq d \end{cases}, \qquad \text{(A.8)}$$

$$\frac{\partial^2 \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{e}, \boldsymbol{\lambda})}{\partial \lambda_d \partial \lambda_m} = \begin{cases} \sum_{\{(k,l):p(k,l)=d\}} [-\frac{b_{kl}}{\lambda_d^2}] & \text{if } d = m \\ 0 & \text{if } d \neq m \end{cases}. \qquad \text{(A.9)}$$

From (A.7)-(A.9), it can be seen that

$$\begin{aligned} H_{11} &< 0, \\ H_{12} &= H_{21} = 0, \\ H_{22} &< 0. \end{aligned}$$

Therefore,

$$\begin{vmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{vmatrix} = H_{11}H_{22} - H_{12}H_{21} > 0.$$

So we conclude that the log-likelihood function is not concave.

A.2 Poisson Model II

The log-likelihood function for the blank scan data is given by

$$\begin{aligned} \phi_2(\boldsymbol{\epsilon}, \lambda) &\triangleq \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}, \lambda) \\ &= \sum_k \sum_{l \in \text{fan}_k} [-\epsilon_k \epsilon_l \lambda c(k,l) + b_{kl} \log(\epsilon_k \epsilon_l \lambda c(k,l))]. \end{aligned} \tag{A.10}$$

The first partial derivatives of $\phi_2(\boldsymbol{\epsilon}, \lambda)$ are

$$\frac{\partial \phi_2(\boldsymbol{\epsilon}, \lambda)}{\partial \epsilon_k} = \sum_{l \in \text{fan}_k} [-\epsilon_l \lambda c(k,l) + \frac{b_{kl}}{\epsilon_k}], \; k = 1, 2, \ldots, I,$$

and

$$\frac{\partial \phi_2(\boldsymbol{\epsilon}, \lambda)}{\partial \lambda} = \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} [-\epsilon_k \epsilon_l + \frac{b_{kl}}{\lambda}], \; k = 1, 2, \ldots, I.$$

And, the second partial derivatives are

$$\begin{aligned}
\frac{\partial^2 \phi_2(\boldsymbol{\epsilon},\lambda)}{\partial \epsilon_k \partial \epsilon_l} &= \begin{cases} -\sum\limits_{l \in \text{fan}_k} \frac{b_{kl}}{\epsilon_k^2} & \text{if } k = l \\ -\lambda c(k,l) & \text{if } l \in \text{fan}_k \\ 0 & \text{otherwise} \end{cases}, \\
\frac{\partial^2 \phi_2(\boldsymbol{\epsilon},\lambda)}{\partial \lambda \partial \epsilon_k} &= -\sum_{l \in \text{fan}_k} \epsilon_l c(k,l), \\
\frac{\partial^2 \phi_2(\boldsymbol{\epsilon},\lambda)}{\partial \lambda^2} &= -\sum_{k=1}^{I} \sum_{l \in \text{fan}_k} \frac{b_{kl}}{\lambda^2}.
\end{aligned} \tag{A.11}$$

From (A.11), we know that

$$H_{11} = -\frac{1}{\epsilon_1^2} \sum_{l \in \text{fan}_1} b_{1l} < 0,$$

$$H_{12} = H_{21} = 0,$$

and

$$H_{22} = -\frac{1}{\epsilon_2^2} \sum_{l \in \text{fan}_2} b_{2l} < 0.$$

Therefore,

$$\begin{vmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{vmatrix} = H_{11}H_{22} - H_{12}H_{21} > 0.$$

So we conclude that $\boldsymbol{H}$ is not negative definite and that the objective function of (**P2**), $\phi_2(\boldsymbol{\epsilon}, \lambda)$, is not concave.

A.3 Poisson Model III

The objective function of the problem (**P3**) is

$$\phi_3(\boldsymbol{\rho}, \lambda) \triangleq \log P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\rho}, \lambda)$$

$$= \sum_k \sum_{l \in \text{fan}_k} [-\rho_k \rho_l \lambda a(k,l) + b_{kl} \log(\rho_k \rho_l \lambda a(k,l))]. \qquad \text{(A.12)}$$

Because Poisson Model III has the exact same functional form as Poisson Model II, the second derivatives of $\phi_3(\boldsymbol{\rho}, \lambda)$ are

$$\begin{aligned}
\frac{\partial^2 \phi_3(\boldsymbol{\rho},\lambda)}{\partial \rho_k \partial \rho_l} &= \begin{cases} -\sum_{l \in \text{fan}_k} \frac{b_{kl}}{\rho_k^2} & \text{if } k = l \\ -\lambda a(k,l) & \text{if } l \in \text{fan}_k \\ 0 & \text{otherwise} \end{cases}, \\
\frac{\partial^2 \phi_3(\boldsymbol{\rho},\lambda)}{\partial \lambda \partial \rho_k} &= -\sum_{l \in \text{fan}_k} \rho_l a(k,l), \\
\frac{\partial^2 \phi_2(\boldsymbol{\rho},\lambda)}{\partial \lambda^2} &= -\sum_{k=1}^{I} \sum_{l \in \text{fan}_k} \frac{b_{kl}}{\lambda^2}.
\end{aligned} \qquad \text{(A.13)}$$

From (A.13), we know that

$$H_{11} = -\frac{1}{\rho_1^2} \sum_{l \in \text{fan}_1} b_{1l} < 0,$$

$$H_{12} = H_{21} = 0,$$

and

$$H_{22} = -\frac{1}{\rho_2^2} \sum_{l \in \text{fan}_2} b_{2l} < 0.$$

Therefore,

$$\begin{vmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{vmatrix} = H_{11} H_{22} - H_{12} H_{21} > 0.$$

So we conclude that $\boldsymbol{H}$ is not negative definite and that the objective function of (**P3**), $\phi_3(\boldsymbol{\rho}, \lambda)$, is not concave.

APPENDIX B
CRAMER-RAO LOWER BOUNDS

To compute the Cramer-Rao lower bound, the Fisher information matrix must be computed. In this appendix, we derive the Fisher information matrix for the sufficiently-parameterized approach associated with each Poisson data model.

B.1 Poisson Model I

For Poisson Model I, the probability density function (PDF) is

$$P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\boldsymbol{\lambda})=\prod_{k=1}^{I}\prod_{l\in\text{fan }k} e^{-\epsilon_k\epsilon_l\lambda_{p(k,l)}}\frac{(\epsilon_k\epsilon_l\lambda_{p(k,l)})^{b_{kl}}}{b_{kl}!}.$$

From (A.3), we know that the second partial derivatives of $\log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\boldsymbol{\lambda})$ are

$$\begin{aligned}
\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\boldsymbol{\lambda})}{\partial\epsilon_k\partial\epsilon_l} &= \begin{cases} -\sum_{l\in\text{fan}_k}\frac{b_{kl}}{\epsilon_k^2} & \text{if } k=l \\ -\lambda_{p(k,l)} & \text{if } l\in\text{fan}_k \\ 0 & \text{otherwise} \end{cases}, \\
\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\boldsymbol{\lambda})}{\partial\lambda_d\partial\epsilon_k} &= -\sum_{l\in\text{fan}_k}[\epsilon_l\delta(p(k,l)-d)], \\
\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\boldsymbol{\lambda})}{\partial\lambda_d\partial\lambda_m} &= \begin{cases} \sum_{\{(k,l):p(k,l)=d\}}\left[-\frac{b_{kl}}{\lambda_d^2}\right] & \text{if } d=m \\ 0 & \text{otherwise} \end{cases}.
\end{aligned} \tag{B.1}$$

The expected value of b_{kl} is $\epsilon_k \epsilon_l \lambda_{p(k,l)}$, so we can calculate the expected values of the second derivatives as follows.

$$
\begin{aligned}
E[\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\boldsymbol{\lambda})}{\partial \epsilon_k \partial \epsilon_l} \mid_{\boldsymbol{b}=\boldsymbol{B}}] &= \begin{cases} -\sum\limits_{l \in \text{fan}_k} \frac{\epsilon_l \lambda_{p(k,l)}}{\epsilon_k} & \text{if } k = l \\ -\lambda_{p(k,l)} & \text{if } l \in \text{fan}_k \\ 0 & \text{otherwise} \end{cases}, \\
E[\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\boldsymbol{\lambda})}{\partial \lambda_d \partial \epsilon_k} \mid_{\boldsymbol{b}=\boldsymbol{B}}] &= -\sum_{l \in \text{fan}_k} [\epsilon_l \delta(p(k,l) - d)], \\
E[\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\boldsymbol{\lambda})}{\partial \lambda_d \partial \lambda_m} \mid_{\boldsymbol{b}=\boldsymbol{B}}] &= \begin{cases} \sum\limits_{\{(k,l):p(k,l)=d\}} [-\frac{\epsilon_k \epsilon_l}{\lambda_d}] & \text{if } d = m \\ 0 & \text{otherwise} \end{cases}.
\end{aligned}
\tag{B.2}
$$

Consider a simplified example, where there are only 4 detectors on the ring (Fig. B.1). Detector pairs (1,3) and (2,4) have the same perpendicular distance from the

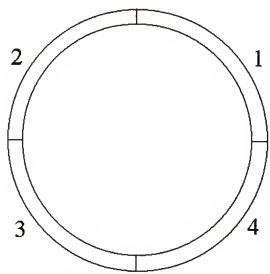


Figure B.1: A simplified example to derive the Fisher information matrix.

origin and therefore have the same λ_1 as the mean of the transmitted photon count. Detector pairs (1,4), (2,1), (3,2), and (4,3) have the same perpendicular distance from the origin and therefore have the same λ_2 as the mean of the transmitted photon count. The data $\{b_{13}, b_{24}, b_{14}, b_{21}, b_{32}, b_{43}\}$ are collected. The problem is to estimate

$\{\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \lambda_1, \lambda_2\}$. Then the Fisher information matrix can be found to be

$$\boldsymbol{I}(\boldsymbol{\epsilon}, \boldsymbol{\lambda}) = \begin{bmatrix} T_1 & \lambda_2 & \lambda_1 & \lambda_2 & \epsilon_3 & \epsilon_2+\epsilon_4 \\ \lambda_2 & T_2 & \lambda_2 & \lambda_1 & \epsilon_4 & \epsilon_1+\epsilon_3 \\ \lambda_1 & \lambda_2 & T_3 & \lambda_2 & \epsilon_1 & \epsilon_2+\epsilon_4 \\ \lambda_2 & \lambda_1 & \lambda_2 & T_4 & \epsilon_2 & \epsilon_1+\epsilon_3 \\ \epsilon_3 & \epsilon_4 & \epsilon_1 & \epsilon_2 & T_5 & 0 \\ \epsilon_2+\epsilon_4 & \epsilon_1+\epsilon_3 & \epsilon_2+\epsilon_4 & \epsilon_1+\epsilon_3 & 0 & T_6 \end{bmatrix},$$

where

$$T_1 = \frac{1}{\epsilon_1}(\epsilon_2\lambda_2 + \epsilon_3\lambda_1 + \epsilon_4\lambda_2),$$

$$T_2 = \frac{1}{\epsilon_2}(\epsilon_1\lambda_2 + \epsilon_3\lambda_2 + \epsilon_4\lambda_1]),$$

$$T_3 = \frac{1}{\epsilon_3}(\epsilon_1\lambda_1 + \epsilon_2\lambda_2 + \epsilon_4\lambda_2),$$

$$T_4 = \frac{1}{\epsilon_4}(\epsilon_1\lambda_2 + \epsilon_2\lambda_1 + \epsilon_3\lambda_2),$$

$$T_5 = \frac{1}{\lambda_1}(\epsilon_1\epsilon_3 + \epsilon_2\epsilon_4),$$

and

$$T_6 = \frac{1}{\lambda_2}(\epsilon_1\epsilon_4 + \epsilon_2\epsilon_1 + \epsilon_3\epsilon_2 + \epsilon_4\epsilon_3).$$

We proved that the determinant of $\boldsymbol{I}(\boldsymbol{\epsilon}, \boldsymbol{\lambda})$ is equal to zero. That is, $\boldsymbol{I}(\boldsymbol{\epsilon}, \boldsymbol{\lambda})$ is non-invertible. The same conclusion can be extended to all higher-dimensional cases.

B.2 Poisson Model II

For Poisson Model II, the PDF is

$$P(\boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}, \lambda) = \prod_{k=1}^{I} \prod_{l \in \text{fan}\, k} e^{-\epsilon_k\epsilon_l\lambda c(k,l)} \frac{(\epsilon_k\epsilon_l\lambda c(k,l))^{b_{kl}}}{b_{kl}!}.$$

From (A.11), we know the second partial derivatives of $\log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\lambda)$ are

$$\begin{aligned}
\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\lambda)}{\partial \epsilon_k \partial \epsilon_l} &= \begin{cases} -\sum\limits_{l\in \text{fan}_k} \frac{b_{kl}}{\epsilon_k^2} & \text{if } k=l \\ -\lambda c(k,l) & \text{if } l \in \text{fan}_k \\ 0 & \text{otherwise} \end{cases}, \\
\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\lambda)}{\partial \lambda \partial \epsilon_k} &= -\sum_{l\in \text{fan}_k} \epsilon_l c(k,l), \\
\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\lambda)}{\partial \lambda^2} &= -\sum_{k=1}^{I} \sum_{l\in \text{fan}_k} \frac{b_{kl}}{\lambda^2}.
\end{aligned} \tag{B.3}$$

The expected values of the second partial derivatives of $\log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\lambda)$ are

$$\begin{aligned}
E\left[\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\lambda)}{\partial \epsilon_k \partial \epsilon_l}\Big|_{\boldsymbol{b}=\boldsymbol{B}}\right] &= \begin{cases} -\sum\limits_{l\in \text{fan}_k} \frac{\epsilon_l \lambda c(k,l)}{\epsilon_k} & \text{if } k=l \\ -\lambda c(k,l) & \text{if } l \in \text{fan}_k \\ 0 & \text{otherwise} \end{cases}, \\
E\left[\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\lambda)}{\partial \lambda \partial \epsilon_k}\Big|_{\boldsymbol{b}=\boldsymbol{B}}\right] &= -\sum_{l\in \text{fan}_k} \epsilon_l c(k,l), \\
E\left[\frac{\partial^2 \log P(\boldsymbol{B}=\boldsymbol{b};\boldsymbol{\epsilon},\lambda)}{\partial \lambda^2}\Big|_{\boldsymbol{b}=\boldsymbol{B}}\right] &= -\sum_{k=1}^{I} \sum_{l\in \text{fan}_k} \frac{\epsilon_k \epsilon_l c(k,l)}{\lambda}.
\end{aligned} \tag{B.4}$$

Consider the example in Fig. B.1 again. There are two possible values for $\{c(k,l)\}$: c_1 for detector pairs (1,3) and (2,4), because these two detector pairs have the same perpendicular distance from the origin; c_2 for detector pairs (1,4), (2,1), (3,2),and (4,3), because these four detector pairs have the same perpendicular distance from the origin. Therefore, the Fisher information matrix is

$$\boldsymbol{I}(\boldsymbol{\epsilon},\lambda) = \begin{bmatrix} S_1 & \lambda c_2 & \lambda c_1 & \lambda c_2 & S_2 \\ \lambda c_2 & S_3 & \lambda c_2 & \lambda c_2 & S_4 \\ \lambda c_1 & \lambda c_2 & S_5 & \lambda c_2 & S_6 \\ \lambda c_2 & \lambda c_1 & \lambda c_2 & S_7 & S_8 \\ S_2 & S_4 & S_6 & S_8 & S_9 \end{bmatrix},$$

where

$$S_1 = \frac{\lambda}{\epsilon_1}(\epsilon_3 c_1 + \epsilon_2 c_2 + \epsilon_4 c_2),$$

$$S_2 = \frac{1}{\lambda}(\epsilon_3 c_1 + \epsilon_2 c_2 + \epsilon_4 c_2),$$

$$S_3 = \frac{\lambda}{\epsilon_2}(\epsilon_1 c_2 + \epsilon_3 c_2 + \epsilon_4 c_1),$$

$$S_4 = \frac{1}{\lambda}(\epsilon_1 c_2 + \epsilon_3 c_2 + \epsilon_4 c_1),$$

$$S_5 = \frac{\lambda}{\epsilon_3}(\epsilon_1 c_1 + \epsilon_2 c_2 + \epsilon_4 c_2),$$

$$S_6 = \frac{1}{\lambda}(\epsilon_1 c_1 + \epsilon_2 c_2 + \epsilon_4 c_2),$$

$$S_7 = \frac{\lambda}{\epsilon_4}(\epsilon_1 c_2 + \epsilon_2 c_1 + \epsilon_3 c_2),$$

$$S_8 = \frac{1}{\lambda}(\epsilon_1 c_2 + \epsilon_2 c_1 + \epsilon_3 c_2),$$

and

$$S_9 = \frac{1}{\lambda}(\epsilon_1\epsilon_3 c_1 + \epsilon_2\epsilon_4 c_1 + \epsilon_1\epsilon_4 c_2 + \epsilon_2\epsilon_1 c_2 + \epsilon_3\epsilon_2 c_2 + \epsilon_4\epsilon_3 c_2).$$

It can be readily proved that, regardless of the values c_1 and c_2, the fifth row of the matrix $\boldsymbol{I}(\boldsymbol{\epsilon}, \lambda)$ can be expressed as a linear combination of the first four rows. Therefore, the Fisher information matrix $\boldsymbol{I}(\boldsymbol{\epsilon}, \lambda)$ has a determinant equal to zero and is non-invertible. The same conclusion can be reached for the general case.

B.3 Poisson Model III

Because Poisson Model III has the exact same functional form as Poisson Model II, we have the same conclusion for Poisson Model III regarding the Fisher information matrix. In other words, the Fisher information matrix is non-invertible.

APPENDIX C
DERIVATION OF THE E STEP OF THE EM ALGORITHMS

In this appendix, we prove that the E step is given by (2.9). Said another way, we determine the conditional expectation of the complete data, $\boldsymbol{X} = [\boldsymbol{B}, \boldsymbol{N}]$, given the incomplete data, $\boldsymbol{B}$, and current estimates of $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$.

C.1 Poisson Model I

Using Bayes' Rule [44], the log-likelihood function of the complete data is given by

$$
\begin{aligned}
L_c(\boldsymbol{b}, \boldsymbol{n}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) &\triangleq \log P(\boldsymbol{B} = \boldsymbol{b}, \boldsymbol{N} = \boldsymbol{n}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) \\
&= \log P(\boldsymbol{B} = \boldsymbol{b} \mid \boldsymbol{N} = \boldsymbol{n}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) P(\boldsymbol{N} = \boldsymbol{n}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) \\
&= \log \prod_{k=1}^{I} \prod_{l \in \text{fan}_k} \binom{n_{kl}}{b_{kl}} (\epsilon_k \epsilon_l)^{b_{kl}} (1 - \epsilon_k \epsilon_l)^{n_{kl} - b_{kl}} \\
&+ \log \prod_{d=d_1}^{d_D} \prod_{\{(k,l):p(k,l)=d\}} e^{-\lambda_d} \frac{(\lambda_d)^{n_{kl}}}{n_{kl}!} \\
&= \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} [\log \binom{n_{kl}}{b_{kl}} + b_{kl} \log(\epsilon_k \epsilon_l) + (n_{kl} - b_{kl}) log(1 - \epsilon_k \epsilon_l)] \\
&+ \sum_{d=d_1}^{d_D} \sum_{\{(k,l):p(k,l)=d\}} [-\lambda_d + n_{kl} \log(\lambda_d) + \log(n_{kl}!)]. \qquad \text{(C.1)}
\end{aligned}
$$

Let $U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m)$ denote the conditional expectation of the log-likelihood function of the complete data given the incomplete data and estimates $\boldsymbol{\epsilon}^m$ and $\boldsymbol{\lambda}^m$. Then,

$$
\begin{aligned}
& U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m) \\
\triangleq\ & E[L_c(\boldsymbol{B}, \boldsymbol{N}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m], \\
=\ & \sum_{k=1}^{I} \sum_{l \in \mathrm{fan}_k} (E[\binom{N_{kl}}{B_{kl}} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m] + \log(\epsilon_k \epsilon_l) E[B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m] \\
+\ & \log(1 - \epsilon_k \epsilon_l) E[N_{kl} - B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m]) \\
+\ & \sum_{d=d_1}^{d_D} \sum_{\{(k,l):p(k,l)=d\}} (-\lambda_d + \log(\lambda_d) E[N_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m] \\
+\ & E[\log(N_{kl}!) \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m]).
\end{aligned}
\tag{C.2}
$$

To maximize $U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m)$ with respect to $\boldsymbol{\epsilon}$, we take into consideration only the terms $\log(\epsilon_k \epsilon_l) E[B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m]$ and $\log(1 - \epsilon_k \epsilon_l) E[N_{kl} - B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m]$ as the remaining terms are independent of $\boldsymbol{\epsilon}$. Similarly, we only need to consider the terms $-\lambda_d$ and $\log(\lambda_d) E[N_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m]$ in order to maximize $U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m)$ with respect to $\boldsymbol{\lambda}$. Let $n_{kl}^m \triangleq E[N_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m]$. Because $E[B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m] = b_{kl}$,

$$
\begin{aligned}
& U(\boldsymbol{\epsilon}, \boldsymbol{\lambda}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m) \\
=\ & \sum_{k=1}^{I} \sum_{l \in \mathrm{fan}_k} [b_{kl} \log(\epsilon_k \epsilon_l) + (n_{kl}^m - b_{kl}) \log(1 - \epsilon_k \epsilon_l)] \\
+\ & \sum_{d=d_1}^{d_D} \sum_{\{(k,l):p(k,l)=d\}} [-\lambda_d + n_{kl}^m \log \lambda_d] + C_1,
\end{aligned}
\tag{C.3}
$$

where C_1 is independent of $\boldsymbol{\epsilon}$ and $\boldsymbol{\lambda}$. To finish the E step, an expression for n_{kl}^m must be obtained.

To find n_{kl}^m, we need to derive the conditional probability of N_{kl} given the incomplete data.

$$\begin{aligned}
& P(N_{kl} = n_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \epsilon, \boldsymbol{\lambda}) \\
= \; & P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) \\
= \; & \frac{P(B_{kl}=b_{kl} \mid N_{kl}=n_{kl}) P(N_{kl}=n_{kl})}{P(B_{kl}=b_{kl})} \\
= \; & \frac{\binom{n_{kl}}{b_{kl}} (\epsilon_k \epsilon_l)^{b_{kl}} (1-\epsilon_k \epsilon_l)^{n_{kl}-b_{kl}} e^{-\lambda_{p(k,l)}} \frac{(\lambda_{p(k,l)})^{n_{kl}}}{n_{kl}!}}{e^{-\epsilon_k \epsilon_l \lambda_{p(k,l)}} \frac{(\epsilon_k \epsilon_l \lambda_{p(k,l)})^{b_{kl}}}{b_{kl}!}} \\
= \; & e^{-(1-\epsilon_k \epsilon_l)\lambda_{p(k,l)}} \frac{[(1-\epsilon_k \epsilon_l)\lambda_{p(k,l)}]^{n_{kl}-b_{kl}}}{(n_{kl}-b_{kl})!}.
\end{aligned} \tag{C.4}$$

The second equality follows from Bayes' Rule. The conditional expectation of N_{kl} is therefore

$$\begin{aligned}
& E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}] \\
= \; & \sum_{n_{kl} \geq b_{kl}} n_{kl} P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \epsilon, \boldsymbol{\lambda}) \\
= \; & \sum_{n_{kl} \geq b_{kl}} (n_{kl} - b_{kl} + b_{kl}) P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) \\
= \; & \sum_{n_{kl} \geq b_{kl}} (n_{kl} - b_{kl}) P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) \\
+ \; & \sum_{n_{kl} \geq b_{kl}} b_{kl} P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}) \\
= \; & \sum_{n'_{kl} \geq 0} n'_{kl} e^{-(1-\epsilon_k \epsilon_l)\lambda_{p(k,l)}} \frac{[(1-\epsilon_k \epsilon_l)\lambda_{p(k,l)}]^{n'_{kl}}}{(n'_{kl})!} \\
+ \; & b_{kl} \sum_{n'_{kl} \geq 0} e^{-(1-\epsilon_k \epsilon_l)\lambda_{p(k,l)}} \frac{[(1-\epsilon_k \epsilon_l)\lambda_{p(k,l)}]^{n'_{kl}}}{(n'_{kl})!}.
\end{aligned} \tag{C.5}$$

where $n'_{kl} = n_{kl} - b_{kl}$. The first summation is equivalent to the mean of a Poisson random variable with parameter $(1-\epsilon_{kl})\lambda_{p(k,l)}$, and the second summation is the sum of the probability over all possible values of n'_{kl}. Therefore,

$$E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \boldsymbol{\lambda}] = (1 - \epsilon_k \epsilon_l)\lambda_{p(k,l)} + b_{kl}. \tag{C.6}$$

Given the previous estimates $\boldsymbol{\epsilon}^m$ and $\boldsymbol{\lambda}^m$,

$$n_{kl}^m \triangleq E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}^m, \boldsymbol{\lambda}^m] = (1 - \epsilon_k^m \epsilon_l^m)\lambda_{p(k,l)}^m + b_{kl}, \tag{C.7}$$

for $k = 1, 2, \ldots, I, l \in \text{fan}_k$.

C.2 Poisson Model II

With $c(k,l)$ known for $k = 1, 2, \ldots, I, l \in \text{fan}_k$, the log-likelihood function of the complete data $\boldsymbol{X} = [\boldsymbol{B}, \boldsymbol{N}]$ is given by

$$
\begin{aligned}
L_c'(\boldsymbol{b}, \boldsymbol{n}; \boldsymbol{\epsilon}, \lambda) &\triangleq \log P(\boldsymbol{B} = \boldsymbol{b}, \boldsymbol{N} = \boldsymbol{n}; \boldsymbol{\epsilon}, \lambda) \\
&= \log P(\boldsymbol{B} = \boldsymbol{b} \mid \boldsymbol{N} = \boldsymbol{n}; \boldsymbol{\epsilon}, \lambda) P(\boldsymbol{N} = \boldsymbol{n}; \boldsymbol{\epsilon}, \lambda) \\
&= \log \prod_{k=1}^{I} \prod_{l \in \text{fan}_k} \binom{n_{kl}}{b_{kl}} (\epsilon_k \epsilon_l)^{b_{kl}} (1 - \epsilon_k \epsilon_l)^{n_{kl} - b_{kl}} \\
&+ \log \prod_{k=1}^{I} \prod_{l \in \text{fan}_k} e^{-\lambda c(k,l)} \frac{(\lambda c(k,l))^{n_{kl}}}{n_{kl}!} \\
&= \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} [\log \binom{n_{kl}}{b_{kl}} + b_{kl} \log(\epsilon_k \epsilon_l) + (n_{kl} - b_{kl}) log(1 - \epsilon_k \epsilon_l)] \\
&+ \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} [-\lambda c(k,l) + n_{kl} \log(\lambda c(k,l)) + \log(n_{kl}!)]. \qquad \text{(C.8)}
\end{aligned}
$$

Let $U'(\boldsymbol{\epsilon}, \lambda; \boldsymbol{\epsilon}^m, \lambda^m)$ denote the conditional expectation of the log-likelihood function of the complete data given the incomplete data and estimates $\boldsymbol{\epsilon}^m$ and λ^m. Then,

$$
\begin{aligned}
& U'(\boldsymbol{\epsilon}, \lambda; \boldsymbol{\epsilon}^m, \lambda^m) \\
\triangleq\ & E[L_c'(\boldsymbol{B}, \boldsymbol{N}; \boldsymbol{\epsilon}, \lambda) \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m], \\
=\ & \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} (E[\binom{N_{kl}}{B_{kl}} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m] + \log(\epsilon_k \epsilon_l) E[B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m] \\
+\ & \log(1 - \epsilon_k \epsilon_l) E[N_{kl} - B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m]) \qquad \text{(C.9)} \\
+\ & \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} (-\lambda c(k,l) + \log(\lambda c(k,l)) E[N_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m] \\
+\ & E[\log(N_{kl}!) \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m]).
\end{aligned}
$$

To maximize $U'(\boldsymbol{\epsilon}, \lambda; \boldsymbol{\epsilon}^m, \lambda^m)$ with respect to $\boldsymbol{\epsilon}$, we take into consideration only the terms $\log(\epsilon_k \epsilon_l) E[B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m]$ and $\log(1 - \epsilon_k \epsilon_l) E[N_{kl} - B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m]$

as the remaining terms are independent of $\boldsymbol{\epsilon}$. Similarly, we only need to consider the terms $-\lambda c(k,l)$ and $\log(\lambda c(k,l))E[N_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m]$ in order to maximize $U'(\boldsymbol{\epsilon}, \lambda; \boldsymbol{\epsilon}^m, \lambda^m)$ with respect to λ. Let $n_{kl}^m \triangleq E[N_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m]$. Because $E[B_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \boldsymbol{\epsilon}^m, \lambda^m] = b_{kl}$,

$$\begin{aligned} & U'(\boldsymbol{\epsilon}, \lambda; \boldsymbol{\epsilon}^m, \lambda^m) \\ = & \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} [b_{kl} \log(\epsilon_k \epsilon_l) + (n_{kl}^m - b_{kl}) \log(1 - \epsilon_k \epsilon_l)] \\ + & \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} [-\lambda c(k,l) + n_{kl}^m \log \lambda c(k,l)] + C_2, \end{aligned} \tag{C.10}$$

where C_2 is independent of $\boldsymbol{\epsilon}$ and λ. To finish the E step, an expression for n_{kl}^m must be obtained.

To find n_{kl}^m, we need to derive the conditional probability of N_{kl} given the incomplete data.

$$\begin{aligned} & P(N_{kl} = n_{kl} \mid \boldsymbol{B} = \boldsymbol{b}; \epsilon, \lambda) \\ = & P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \lambda) \\ = & \frac{P(B_{kl}=b_{kl} \mid N_{kl}=n_{kl}) P(N_{kl}=n_{kl})}{P(B_{kl}=b_{kl})} \\ = & \frac{\binom{n_{kl}}{b_{kl}} (\epsilon_k \epsilon_l)^{b_{kl}} (1-\epsilon_k \epsilon_l)^{n_{kl}-b_{kl}} e^{-\lambda c(k,l)} \frac{(\lambda c(k,l))^{n_{kl}}}{n_{kl}!}}{e^{-\epsilon_k \epsilon_l \lambda c(k,l)} \frac{(\epsilon_k \epsilon_l \lambda c(k,l))^{b_{kl}}}{b_{kl}!}} \\ = & e^{-(1-\epsilon_k \epsilon_l)\lambda c(k,l)} \frac{[(1-\epsilon_k \epsilon_l)\lambda c(k,l)]^{n_{kl}-b_{kl}}}{(n_{kl}-b_{kl})!}. \end{aligned} \tag{C.11}$$

The second equality follows from Bayes' Rule. The conditional expectation of N_{kl} is therefore

$$\begin{aligned}
& E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \lambda] \\
= & \sum_{n_{kl} \geq b_{kl}} n_{kl} P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \epsilon, \lambda) \\
= & \sum_{n_{kl} \geq b_{kl}} (n_{kl} - b_{kl} + b_{kl}) P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \lambda) \\
= & \sum_{n_{kl} \geq b_{kl}} (n_{kl} - b_{kl}) P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \lambda) \\
+ & \sum_{n_{kl} \geq b_{kl}} b_{kl} P(N_{kl} = n_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \lambda) \\
= & \sum_{n'_{kl} \geq 0} n'_{kl} e^{-(1-\epsilon_k \epsilon_l)\lambda c(k,l)} \frac{[(1-\epsilon_k \epsilon_l)\lambda c(k,l)]^{n'_{kl}}}{(n'_{kl})!} \\
+ & b_{kl} \sum_{n'_{kl} \geq 0} e^{-(1-\epsilon_k \epsilon_l)\lambda c(k,l)} \frac{[(1-\epsilon_k \epsilon_l)\lambda c(k,l)]^{n'_{kl}}}{(n'_{kl})!}.
\end{aligned} \tag{C.12}$$

where $n'_{kl} = n_{kl} - b_{kl}$. The first summation is equivalent to the mean of a Poisson random variable with parameter $(1 - \epsilon_{kl})\lambda c(k,l)$, and the second summation is the sum of the probability over all possible values of n'_{kl}. Therefore,

$$E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}, \lambda] = (1 - \epsilon_k \epsilon_l)\lambda c(k,l) + b_{kl}. \tag{C.13}$$

Given the previous estimates $\boldsymbol{\epsilon}^m$ and λ^m,

$$n_{kl}^m \triangleq E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\epsilon}^m, \lambda^m] = (1 - \epsilon_k^m \epsilon_l^m)\lambda^m c(k,l) + b_{kl}, \tag{C.14}$$

for $k = 1, 2, \ldots, I, l \in \text{fan}_k$.

C.3 Poisson Model III

Because Poisson Model III has the exact same functional form as Poisson Model II, the E step associated with Poisson Model III follows from the same derivation as that associated with Poisson Model II. Therefore, the conditional expectation of the log-likelihood function of the complete data $\boldsymbol{X} = [\boldsymbol{B}, \boldsymbol{N}]$ given the incomplete data

and current estimates $\boldsymbol{\rho}^m$ and λ^m is

$$\begin{aligned} & U''(\boldsymbol{\rho}, \lambda; \boldsymbol{\rho}^m, \lambda^m) \\ = & \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} [b_{kl} \log(\rho_k \rho_l) + (n_{kl}^m - b_{kl}) \log(1 - \rho_k \rho_l)] \\ + & \sum_{k=1}^{I} \sum_{l \in \text{fan}_k} [-\lambda a(k,l) + n_{kl}^m \log \lambda a(k,l)] + C_3, \end{aligned} \tag{C.15}$$

where C_3 is independent of $\boldsymbol{\rho}$ and λ, and

$$n_{kl}^m \triangleq E[N_{kl} \mid B_{kl} = b_{kl}; \boldsymbol{\rho}^m, \lambda^m] = (1 - \rho_k^m \rho_l^m) \lambda^m a(k,l) + b_{kl}, \tag{C.16}$$

for $k = 1, 2, \ldots, I, l \in \text{fan}_k$.

REFERENCES

[1] K. Lange, and R. Carson, "EM Reconstruction Algorithms for Emission and Transmission Tomography," *Journal of Computer Assisted Tomography*, vol. 8, no. 2, pp. 306-316, Apr. 1984.

[2] H. Hiriyannaiah, "X-Ray Computed Tomography for Medical Imaging," *IEEE Signal Processing Magazine*, pp. 42-59, March, 1997.

[3] K. Lee, K. Jung, Y. Yoon, I. Jeung, and K. Jeong, "The Development of Planar Imaging Technique for Spray Characterization," *Conference on Lasers and Electro-Optics*, vol. 2, pp. 163-164, Pacific Rim, 1999.

[4] W. Klein, H. Barrett, I. Pang, D. Patton, M. Rogulski, J. Sain, and W. Smith, "FASTSPECT: Electrical and Mechanical Design of a High-Resolution Dynamic SPECT Imager," *IEEE Medical Imaging Conference*, vol. 2, pp. 931-933, Oct. 1995.

[5] E. Garcia, T. Faber, J. Galt, C. Cooke, and R. Folks, "Advances in Nuclear Emission PET and SPECT Imaging," *IEEE Engineering in Medicine and Biology Magazine*, vol. 19, no. 5. pp. 21-33, Sep.-Oct. 2000.

[6] K. Iwata, M. Wu, and B. Hasegawa, "Design of Combined X-Ray CT and SPECT Systems for Small Animals," *Nuclear Science Symposium*, vol. 3, pp. 1608-1612, 1999.

[7] J. Ollinger, and J. Fessler, "Positron-Emission Tomography," *IEEE Signal Processing Magazine*, vol. 14, no. 1, pp. 43-55, Jan. 1997.

[8] C. Wu, "Attenuation Correction Methods for Positron Emission Tomography," Ph. D. Dissertation, University of Florida, pp. 4-5, May, 1998.

[9] M. Defrise, D. Townsend, D. Bailey, A. Geissbuhler, C. Michel, and T. Jones, "A Normalization Technique for 3D PET Data," *Phy. Med. Biol.*, vol. 36, no. 7, pp. 939-952, 1991.

[10] E. Hoffman, S. Huang, M. Phelps, and D. Kuhl, "Quantization in Positron-Emission Computed Tomography," *Journal of Computer Assisted Tomography*, vol. 5, pp. 391-400, 1981.

[11] M. Daube-Witherspoon and R. Carson, "Unified Deadtime Correction Model for PET," *IEEE Trans. Medical Imaging*, vol. 10, pp. 267-275, 1991.

[12] D. Bailey, D. Townsend, P. Kinahan, S. Grootoonk, and T. Jones, "An Investigation of Factors Affecting Detector and Geometric Correction in Normalization of 3-D PET Data," *IEEE Trans. on Nuclear Science*, vol. 43, no. 6, pp. 3300-3307, Dec. 1996.

[13] S. Derenzo, "Mathematical Removal of Positron Range Blurring in High Resolution Tomography," *IEEE Trans. Nuclear Science*, vol. 33, pp. 565-569, 1986.

[14] L. Shepp, and Y. Vardi, "Maximum Likelihood Reconstruction for Emission Tomography," *IEEE Trans. on Medical Imaging*, vol. I-1, no. 2, pp. 113-122, Oct. 1982.

[15] G. Grimmett, and D. Stirzaker, *Probability and Random Processes*, pp. 255, New York: Oxford 2001.

[16] E. Hoffman, T. Guerrero, G. Germano, W. Digby, and M. Dahlbom, "PET System Calibrations and Corrections for Quantitative and Spatially Accurate Images," *IEEE Trans. on Nuclear Science*, vol. 36, no. 1, pp. 1108-1112, Feb. 1989.

[17] N. Ferreira, R. Trébossen, M. Grégoire, and B. Bendriem, "Influence of Malfunctioning Block Detectors on the Calculation of Single Detector Efficiencies in PET," *IEEE Trans. on Nuclear Science*, vol. 46, no. 4, pp. 1062-1069, Aug. 1999.

[18] Y. Pawitan, S. Kohlmyer, T. Lewellen, and F. O'Sullivan, "PET System Calibration and Attenuation Correction," *IEEE Trans. on Nuclear Science*, vol. 44, no. 3, pp. 1249-1253, June 1997.

[19] F. Hermansen, T. Spinks, P. Camici, and A. Lammertsma, "Calculation of Single Detector Efficiencies and Extension of the Normalization Sinogram in PET," *Phys. Med. Biol.*, vol. 42, no. 6, pp. 1143-1154, June 1997.

[20] J. Ollinger, "Detector Efficiency and Compton Scatter in Fully 3D PET," *IEEE Trans. on Nuclear Science*, vol. 42, no. 4, pp. 1168-1173, Aug. 1995.

[21] R. Badawi, M, Lodge, and P. Marsden, "Algorithms for Calculating Detector Efficiency Normalization Coefficients for True Coincidences in 3D PET," *Phy. Med. Biol.*, 43. pp. 189-205, 1998.

[22] D. Chesler, and C. Sterns, "Calibration of Detector Sensitivity in Positron Cameras," *IEEE Trans. on Medical Science*, vol. 37, no. 2, pp. 768-772, 1990.

[23] M. Casey, and E. Hoffman, "Quantitation in Positron Emission Computed Tomography: 7. A Technique to Reduce Noise in Accidental Coincidence Measurements and Coincidence Efficiency Calibration," *Journal of Computer Assisted Tomography*, vol. 10, no. 5, pp. 845-850, Sep. 1986.

[24] M. Casey, H. Gadagkar, and D. Newport, "A Component Based Method for Normalization in Volume PET," *Proc. 3rd International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, Aix-Ics-Bains, pp. 67-71, 1995.

[25] Z. Liang, "Detector Response Restoration in Image Reconstruction of High Resolution Positron Emission Tomography," *IEEE Trans. on Medical Imaging*, vol. 13, no. 2, pp. 314-321, June 1994.

[26] K. Wienhard, L. Eriksson, S. Grootoonk, M. Casey, U. Pietrzyk, and W. Heiss, "Performance Evaluation of the Positron Scanner ECAT EXACT," *Journal of Computer Assisted Tomography*, vol. 16, no. 5, pp. 804-813, 1992.

[27] Z. Cho, S. Juh, R. Friedenberg, W. Bunney, M. Buchsbaum, and E. Wong, "A New Approach to Very High Resolution Mini-Brain PET Using a Small Number of Large Detectors," *IEEE Trans. on Nuclear Science*, vol. 37, no. 2. pp. 842-851, April 1990.

[28] R. Huesman, E. Salmeron, and J. Baker, "Compensation for Crystal Penetration in High Resolution Positron Tomography," *IEEE Trans. on Nuclear Science*, vol. 36, no. 1, pp. 1100-1107, Feb. 1989.

[29] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, pp. 157-198, Englewood Cliffs, NJ: Prentice-Hall 1993.

[30] A. Dempster, N. Laird, and D. Rubin, "Maximum Likelihood from Incomplete Data via the EM algorithm," *Ann. Roy. Stat. Soc.*, vol. 39, no. 1, pp. 1-38, Dec. 1977.

[31] M. Feder, and E. Weistein, "Parameter Estimation of Superimposed Signals Using the EM Algorithm," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477-489, Apr. 1988.

[32] M. Bazaraa, H. Sherali, and C. Shetty, *Nonlinear Programming: Theory and Algorithms*, pp. 268-275, New York: Wiley 1979.

[33] T. Abatzoglou, and B. O'Donnell, "Minimization by Coordinate Descent," *J. of Opt. Theory and Applications*, vol. 36, no. 2, pp. 163-174, Feb. 1982.

[34] Z. Luo, and P. Tseng, "On the Convergence of the Coordinate Descent Method for Convex Differentiable Minimization," *J. of Opt. Theory and Applications*, vol. 72, no. 1, pp. 7-35, Jan. 1992.

[35] P. Stoica, and T. Marzetta, "Parameter Estimation Problems with Singular Information Matrices," *IEEE Trans. on Signal Processing*, vol. 49, no. 1, pp. 87-90, Jan. 2001.

[36] R. Carroll, "PET Reconstruction Algorithm," Master Thesis, Department of Mathematics, University of Florida, 1998.

[37] Y. Vardi, L. Shepp, and Y. Kaufman, "A Statistical Model for Positron Emission Tomography"' *Journal of the American Statistical Association*, vol. 80, no. 389, pp. 8-20, March 1985.

[38] M. Madsen, "PET Attenuation Correction Using Mean Attenuation Coefficients: a Simulation Study," *IEEE Trans. on Nuclear Science*, vol. 46, no. 6, pp. 2172-2176, Dec. 1999.

[39] A. Papenfuss, G. O'Keefe, and A. Scott, "Segmented Attenuation Correction in Whole Body PET Using Neigh EM Clustering," *IEEE Nuclear Science Symposium Conference Record*, vol. 2, no. 2, pp. 13/78-13/81, 2000.

[40] A. De Pierro, "On the Relation Between the ISRA and the EM Algorithm for Positron Emission Tomography," *IEEE Trans. on Medical Imaging*, vol. 12, no. 2, pp. 328-333, June, 1993.

[41] H. Hudson, and R. Larkin, "Accelerated Image Reconstruction Using Ordered Subsets of Projection Data," *IEEE Trans. on Medical Imaging*, vol. 13, no. 4, pp. 601-609, Dec. 1994.

[42] K. Choi, Personal Communication, University of Florida.

[43] B. Noble, and J. Daniel, *Applied Linear Algebra*, pp. 427, Englewood Cliffs, NJ: Prentice-Hall 1977.

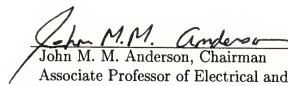
[44] Y. Viniotis, *Probability and Random Processes*, pp. 38, New York: McGraw-Hill 1998.

BIOGRAPHICAL SKETCH

I was born in Kaohsiung, a big modern city in Taiwan. After I earned my B.S. degree from the Electronics Engineering Department of National Jiao Tong University in Hsinchu, I was admitted by the graduate school of University of Florida and studied for Master and Ph. D. degrees in the Department of Electrical and Computer Engineering. My research focuses on digital signal processing. Since I joined the research group led by my advisor, Dr. Anderson, I had been working in the Statistical Signal Processing Lab and doing research on medical image processing.

Apart from academics, I enjoy reading Chinese literature, listening to music, and playing basketball.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



John M. M. Anderson, Chairman
Associate Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William W. Edmonson
Associate Professor of Electrical Engineering, Hampton University

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jian Li
Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Panagote M. Pardalos
Professor of Industrial and Systems Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 2001

Pramod P. Khargonekar
Dean, College of Engineering

Winfred M. Phillips
Dean, Graduate School